Schultz, J.
10/673063
Seq ID 3

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 24, 2005, 10:20:44 ; Search time 1762 Seconds
(without alignments)
1210.005 Million cell updates/sec

Title: US-10-673-063-3_COPY_900_943
Perfect score: 44
Sequence: 1 gcgggtcccgttccttctta.....ccggtcgcggttattaagaa 44

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 1839042

Minimum DB seq length: 0
Maximum DB seq length: 50

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 100 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
c	1	16.2	36.8	32	6	AX468571	AX468571 Sequence
c	2	16	36.4	27	6	AX338903	AX338903 Sequence
	3	16	36.4	27	6	AX338904	AX338904 Sequence
	4	16	36.4	38	6	A09534	A09534 Oligonucleo
	5	16	36.4	38	6	A09567	A09567 Oligonucleo
c	6	16	36.4	50	12	SYNRNA3C	M19550 Synthetic B
c	7	15.8	35.9	38	6	A17381	A17381 Nucleotide
	8	15.8	35.9	41	6	AX515653	AX515653 Sequence
	9	15.8	35.9	41	6	AX518248	AX518248 Sequence
	10	15.8	35.9	50	3	DROPRDM1	K03518 D.melanogas
	11	15.2	34.5	33	6	AR365669	AR365669 Sequence
	12	15.2	34.5	37	6	AX147162	AX147162 Sequence
c	13	15.2	34.5	47	6	AX081593	AX081593 Sequence
c	14	15.2	34.5	47	6	AX374774	AX374774 Sequence
c	15	15	34.1	31	6	AX248444	AX248444 Sequence
	16	15	34.1	50	10	MMU41925	U41925 Mus musculu
c	17	14.8	33.6	31	6	AX582316	AX582316 Sequence
c	18	14.8	33.6	33	6	I03806	I03806 Sequence 9
	19	14.8	33.6	36	6	AR429930	AR429930 Sequence
	20	14.8	33.6	36	6	AX099684	AX099684 Sequence
c	21	14.8	33.6	37	6	I03802	I03802 Sequence 5
c	22	14.8	33.6	47	6	AR288505	AR288505 Sequence
c	23	14.8	33.6	47	6	AR289990	AR289990 Sequence
c	24	14.8	33.6	50	6	AR032973	AR032973 Sequence
c	25	14.8	33.6	50	6	I29713	I29713 Sequence 58
c	26	14.8	33.6	50	6	I91387	I91387 Sequence 58
c	27	14.8	33.6	50	6	AR209637	AR209637 Sequence
	28	14.6	33.2	32	6	CQ868809	CQ868809 Sequence
	29	14.6	33.2	33	6	AR365650	AR365650 Sequence
c	30	14.6	33.2	34	6	AR365637	AR365637 Sequence
c	31	14.6	33.2	35	6	AX739899	AX739899 Sequence
	32	14.6	33.2	39	6	AR365666	AR365666 Sequence
c	33	14.6	33.2	45	6	AX642259	AX642259 Sequence
	34	14.6	33.2	46	6	I13724	I13724 Sequence 30
	35	14.6	33.2	47	6	AR291680	AR291680 Sequence
	36	14.4	32.7	20	6	AX462584	AX462584 Sequence
	37	14.4	32.7	32	6	BD061913	BD061913 Antigenic
	38	14.4	32.7	34	6	AX574343	AX574343 Sequence
	39	14.4	32.7	40	6	BD235752	BD235752 Strengthe
	40	14.4	32.7	40	6	BD235803	BD235803 Method of
	41	14.4	32.7	47	6	AR291235	AR291235 Sequence
c	42	14.4	32.7	50	5	XELRGT3	M13248 X.laevis ri
c	43	14.4	32.7	50	6	AX057065	AX057065 Sequence
c	44	14.2	32.3	20	6	AR316182	AR316182 Sequence
c	45	14.2	32.3	20	6	AR337058	AR337058 Sequence
c	46	14.2	32.3	21	6	AR529926	AR529926 Sequence
c	47	14.2	32.3	21	6	AX095951	AX095951 Sequence
c	48	14.2	32.3	25	6	AX304716	AX304716 Sequence
c	49	14.2	32.3	25	6	AX615111	AX615111 Sequence
c	50	14.2	32.3	33	6	AX280063	AX280063 Sequence
	51	14.2	32.3	36	6	I18300	I18300 Sequence 3
c	52	14.2	32.3	38	6	AR089801	AR089801 Sequence
	53	14.2	32.3	48	8	AJ835268	AJ835268 Arabidops
	54	14.2	32.3	48	8	AJ835276	AJ835276 Arabidops
c	55	14.2	32.3	49	6	AR239848	AR239848 Sequence
c	56	14.2	32.3	49	6	AX279650	AX279650 Sequence
	57	14.2	32.3	50	6	AX190235	AX190235 Sequence
	58	14.2	32.3	50	10	MMU41975	U41975 Mus musculu
	59	14	31.8	24	6	A41608	A41608 Sequence 17
	60	14	31.8	24	6	A95670	A95670 Sequence 17
	61	14	31.8	24	6	A95712	A95712 Sequence 17
	62	14	31.8	24	6	A95754	A95754 Sequence 17
	63	14	31.8	24	6	A95796	A95796 Sequence 17
	64	14	31.8	24	6	AR116287	AR116287 Sequence
	65	14	31.8	32	6	AR534294	AR534294 Sequence
	66	14	31.8	32	6	AR544580	AR544580 Sequence
	67	14	31.8	35	6	AR534275	AR534275 Sequence
	68	14	31.8	35	6	AR544561	AR544561 Sequence
c	69	14	31.8	38	6	A86752	A86752 Sequence 12
c	70	14	31.8	38	6	BD062648	BD062648 Interacti
c	71	14	31.8	41	6	AR305163	AR305163 Sequence
c	72	14	31.8	41	6	AR309267	AR309267 Sequence
c	73	14	31.8	41	6	BD106074	BD106074 Novel LDL
	74	14	31.8	45	6	AX763470	AX763470 Sequence
	75	14	31.8	47	6	AR032638	AR032638 Sequence
	76	14	31.8	47	6	I29378	I29378 Sequence 25
	77	14	31.8	47	6	I91052	I91052 Sequence 25
	78	14	31.8	47	6	AR209302	AR209302 Sequence
	79	14	31.8	47	6	AR289255	AR289255 Sequence
c	80	13.8	31.4	24	6	AX445527	AX445527 Sequence
c	81	13.8	31.4	25	6	AX115688	AX115688 Sequence
	82	13.8	31.4	26	6	I21917	I21917 Sequence 2
c	83	13.8	31.4	26	6	AX134798	AX134798 Sequence
c	84	13.8	31.4	26	6	AX137764	AX137764 Sequence
	85	13.8	31.4	29	6	AR160337	AR160337 Sequence
	86	13.8	31.4	34	6	AR157567	AR157567 Sequence
	87	13.8	31.4	34	6	AR212616	AR212616 Sequence
	88	13.8	31.4	34	6	AR430014	AR430014 Sequence
	89	13.8	31.4	34	6	AR533433	AR533433 Sequence
c	90	13.8	31.4	35	6	I03808	I03808 Sequence 11
c	91	13.8	31.4	37	6	AX800567	AX800567 Sequence
c	92	13.8	31.4	38	6	AR272268	AR272268 Sequence

c 93	13.8	31.4	42	6	CQ767048	CQ767048 Sequence
94	13.8	31.4	45	6	E05105	E05105 PCR primer
95	13.8	31.4	45	11	AL834120	AL834120 Arabidops
c 96	13.8	31.4	47	6	AR382719	AR382719 Sequence
97	13.8	31.4	48	6	BD263286	BD263286 Compositi
98	13.8	31.4	48	6	BD263298	BD263298 Compositi
99	13.8	31.4	48	6	BD263303	BD263303 Compositi
100	13.8	31.4	48	6	BD263304	BD263304 Compositi

ALIGNMENTS

RESULT 1
AX468571/c
LOCUS AX468571 32 bp DNA linear PAT 16-JUL-2002
DEFINITION Sequence 12 from Patent WO0238745.
ACCESSION AX468571
VERSION AX468571.1 GI:21901398
KEYWORDS .
SOURCE synthetic construct
ORGANISM synthetic construct
other sequences; artificial sequences.
REFERENCE 1
AUTHORS Meskiene,I., Hirt,H. and Jonak,C.
TITLE Regulation of mitogen-activated protein kinase (mapk)
JOURNAL Patent: WO 0238745-A 12 16-MAY-2002;
Oesterreichisches Forschungszentrum Seibersdorf GmbH (AT)
FEATURES Location/Qualifiers
source 1. .32
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="primer"
ORIGIN

Query Match 36.8%; Score 16.2; DB 6; Length 32;
Best Local Similarity 85.7%; Pred. No. 8e+04;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
Qy         23 AACCGGTCGCGGTTATTAAGA 43
              |||  |||||||||||| |||
Db         31 AACATGTCGCGGTTATGAAGA 11
```

RESULT 2
AX338903/c
LOCUS AX338903 27 bp DNA linear PAT 09-JAN-2002
DEFINITION Sequence 8 from Patent WO0185971.
ACCESSION AX338903
VERSION AX338903.1 GI:18129070
KEYWORDS .
SOURCE synthetic construct
ORGANISM synthetic construct
other sequences; artificial sequences.
REFERENCE 1
AUTHORS Alberte,R.S. and Smith,R.D.
TITLE Transgenic plants incorporating traits of Zostera marina
JOURNAL Patent: WO 0185971-A 8 15-NOV-2001;
Phycogen, Inc. (US)
FEATURES Location/Qualifiers
source 1. .27
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="Primer"
ORIGIN

Query Match 36.4%; Score 16; DB 6; Length 27;
Best Local Similarity 79.2%; Pred. No. 9.9e+04;
Matches 19; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

```
Qy         20 AATAACCGGTCGCGGTTATTAAGA 43
              |||||| |||| ||||| | |||
Db         25 AATAACTTGTCGGGGTTAATCAGA 2
```

RESULT 3
AX338904
LOCUS AX338904 27 bp DNA linear PAT 09-JAN-2002
DEFINITION Sequence 9 from Patent WO0185971.
ACCESSION AX338904
VERSION AX338904.1 GI:18129071
KEYWORDS .
SOURCE synthetic construct
ORGANISM synthetic construct
other sequences; artificial sequences.
REFERENCE 1
AUTHORS Alberte,R.S. and Smith,R.D.
TITLE Transgenic plants incorporating traits of Zostera marina
JOURNAL Patent: WO 0185971-A 9 15-NOV-2001;
Phycogen, Inc. (US)
FEATURES Location/Qualifiers
source 1. .27
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="Primer"
ORIGIN

Query Match 36.4%; Score 16; DB 6; Length 27;
Best Local Similarity 79.2%; Pred. No. 9.9e+04;
Matches 19; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

```
Qy         20 AATAACCGGTCGCGGTTATTAAGA 43
              |||||| |||| ||||| | |||
Db          3 AATAACTTGTCGGGGTTAATCAGA 26
```

RESULT 4
A09534
LOCUS A09534 38 bp DNA linear PAT 02-SEP-2002
DEFINITION Oligonucleotide.
ACCESSION A09534
VERSION A09534.1 GI:411963
KEYWORDS .
SOURCE synthetic construct
ORGANISM synthetic construct
other sequences; artificial sequences.
REFERENCE 1 (bases 1 to 38)
AUTHORS Fleer,R., Fukuhara,H. and Yeh,P.
TITLE Method for the microbiological preparation of human serum albumin
and other heterologous proteins from a yeast
JOURNAL Patent: EP 0361991-A 5 04-APR-1990;
RHONE-POULENC SANTE
FEATURES Location/Qualifiers
source 1. .38
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
ORIGIN

Query Match 36.4%; Score 16; DB 6; Length 38;
Best Local Similarity 68.8%; Pred. No. 9.8e+04;
Matches 22; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

```
Qy         11 TTCCTTCTTAATAACCGGTCGCGGTTATTAAG 42
              ||| ||||  |||| ||| ||||    ||| |
Db          7 TTCTTTCTCGATAAGCGGCCGCGTGCTTTATG 38
```

RESULT 5
A09567
LOCUS A09567 38 bp DNA linear PAT 02-SEP-2002
DEFINITION Oligonucleotide.
ACCESSION A09567

```
VERSION     A09567.1  GI:411995
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            other sequences; artificial sequences.
REFERENCE   1  (bases 1 to 38)
  AUTHORS   Fleer,R., Fukuhara,H. and Yeh,P.
  TITLE     Method for the microbiological preparation of human serum albumin
            and other heterologous proteins from a yeast
  JOURNAL   Patent: EP 0361991-A 39 04-APR-1990;
            RHONE-POULENC SANTE
FEATURES             Location/Qualifiers
     source          1. .38
                     /organism="synthetic construct"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32630"
ORIGIN

  Query Match           36.4%;  Score 16;  DB 6;  Length 38;
  Best Local Similarity   68.8%;  Pred. No. 9.8e+04;
  Matches   22;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy         11 TTCCTTCTTAATAACCGGTCGCGGTTATTAAG 42
              ||| ||||  |||| ||| ||||    ||| |
Db          7 TTCTTTCTCGATAAGCGGCCGCGTGCTTTATG 38


RESULT 6
SYNRNA3C/c
LOCUS       SYNRNA3C                  50 bp ss-RNA     linear   SYN 27-APR-1993
DEFINITION  Synthetic Brome mosaic virus (BMV) RNA 3' end/CAT gene from
            pB3CA42, partial cds.
ACCESSION   M19550
VERSION     M19550.1  GI:209273
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            other sequences; artificial sequences.
REFERENCE   1  (bases 1 to 50)
  AUTHORS   French,R., Janda,M. and Ahlquist,P.G.
  TITLE     Bacterial gene inserted in an engineered RNA virus: Efficient
            expression in monocotyledonous plant cells
  JOURNAL   Science 231, 1294-1297 (1986)
COMMENT     Original source text: Brome mosaic virus and plasmid pB3CA42 RNA.
FEATURES             Location/Qualifiers
     source          1. .50
                     /organism="synthetic construct"
                     /mol_type="genomic RNA"
                     /db_xref="taxon:32630"
     CDS             10. .>50
                     /note="coat protein"
                     /codon_start=1
                     /transl_table=11
                     /protein_id="AAA72637.1"
                     /db_xref="GI:554572"
                     /translation="MSTRFSGAKEAKM"
ORIGIN

  Query Match           36.4%;  Score 16;  DB 12;  Length 50;
  Best Local Similarity   62.5%;  Pred. No. 9.8e+04;
  Matches   25;  Conservative    0;  Mismatches   15;  Indels    0;  Gaps    0;

Qy          4 GGTCCCGTTCCTTCTTAATAACCGGTCGCGGTTATTAAGA 43
              | | || |  || || || | |  ||||   ||||||| |
Db         41 GCTTCCTTAGCTCCTGAAAATCTCGTCGACATTATTAATA 2


RESULT 7
A17381/c
LOCUS       A17381                    38 bp    DNA     linear   PAT 27-APR-1994
DEFINITION  Nucleotide sequence 6 from patent number EP0481502.
ACCESSION   A17381
VERSION     A17381.1  GI:513877
KEYWORDS    .
SOURCE      unidentified
  ORGANISM  unidentified
            unclassified.
REFERENCE   1  (bases 1 to 38)
  AUTHORS   Weidle,U.H. and Kaluza,B.
  TITLE     Process for the production of chimaeric antibodies
  JOURNAL   Patent: EP 0481502-A 6 22-APR-1992;
            BOEHRINGER MANNHEIM GMBH
FEATURES             Location/Qualifiers
     source          1. .38
                     /organism="unidentified"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32644"
ORIGIN

  Query Match           35.9%;  Score 15.8;  DB 6;  Length 38;
  Best Local Similarity   65.7%;  Pred. No. 1.2e+05;
  Matches   23;  Conservative    0;  Mismatches   12;  Indels    0;  Gaps    0;

Qy          1 GCGGGTCCCGTTCCTTCTTAATAACCGGTCGCGGT 35
              ||   ||| | || || ||  ||| ||| |||| |
Db         36 GCCTCTCCAGGTCTTTATTTTTAAGCGGCCGCGCT 2


RESULT 8
AX515653
LOCUS       AX515653                  41 bp    DNA     linear   PAT 05-OCT-2002
DEFINITION  Sequence 1851 from Patent WO02052044.
ACCESSION   AX515653
VERSION     AX515653.1  GI:23562954
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Nakamura,Y., Sekine,A., Iida,A. and Saito,S.
  TITLE     Detection of genetic polymorphisms
  JOURNAL   Patent: WO 02052044-A 1851 04-JUL-2002;
            Riken (JP)
FEATURES             Location/Qualifiers
     source          1. .41
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

  Query Match           35.9%;  Score 15.8;  DB 6;  Length 41;
  Best Local Similarity   81.0%;  Pred. No. 1.2e+05;
  Matches   17;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          5 GTCCCGTTCCTTCTTAATAAC 25
              |||||:||||||| | ||| |
Db         16 GTCCCRTTCCTTCATTATATC 36


RESULT 9
AX518248
LOCUS       AX518248                  41 bp    DNA     linear   PAT 05-OCT-2002
DEFINITION  Sequence 4446 from Patent WO02052044.
ACCESSION   AX518248
VERSION     AX518248.1  GI:23567646
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Nakamura,Y., Sekine,A., Iida,A. and Saito,S.
  TITLE     Detection of genetic polymorphisms
  JOURNAL   Patent: WO 02052044-A 4446 04-JUL-2002;
            Riken (JP)
```

```
FEATURES             Location/Qualifiers
     source          1. .41
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

  Query Match             35.9%;  Score 15.8;  DB 6;  Length 41;
  Best Local Similarity   81.0%;  Pred. No. 1.2e+05;
  Matches   17;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          5 GTCCCGTTCCTTCTTAATAAC 25
              |||||:|||||||  | ||| |
Db         16 GTCCCRTTCCTTCATTATATC 36


RESULT 10
DROPRDM1
LOCUS       DROPRDM1                  50 bp    DNA     linear   INV 26-APR-1993
DEFINITION  D.melanogaster paired (prd) gene with a 1.1 kb insertion in exon 2,
            5' recombination site.
ACCESSION   K03518
VERSION     K03518.1  GI:158175
KEYWORDS    insertion sequence; segmentation gene.
SEGMENT     1 of 2
SOURCE      Drosophila melanogaster (fruit fly)
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 50)
  AUTHORS   Frigerio,G., Burri,M., Bopp,D., Baumgartner,S. and Noll,M.
  TITLE     Structure of the segmentation gene paired and the Drosophila PRD
            gene set as part of a gene network
  JOURNAL   Cell 47 (5), 735-746 (1986)
  MEDLINE   87051745
   PUBMED   2877746
COMMENT     Original source text: D.melanogaster (X-ray induced mutant strain
            PRD-2.45.17) DNA, library of C.Nuesslein-Volhard.
            The insert in mutant Drosophila exon 2, described in [1], causes
            the prd mRNA to increase in size by 1.1 kb.  The process also
            deletes five base pairs from the wild-type prd gene, starting at
            position 1065.
FEATURES             Location/Qualifiers
     source          1. .50
                     /organism="Drosophila melanogaster"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:7227"
     CDS             <1. .>50
                     /note="pseudoPRD cds"
                     /pseudo
                     /codon_start=1
ORIGIN      Chromosome 2 band 33C1,2.

  Query Match             35.9%;  Score 15.8;  DB 3;  Length 50;
  Best Local Similarity   74.1%;  Pred. No. 1.2e+05;
  Matches   20;  Conservative    0;  Mismatches    7;  Indels    0;  Gaps    0;

Qy         12 TCCTTCTTAATAACCGGTCGCGGTTAT 38
              | | ||   ||||||||||||| || ||
Db         17 TGCATCCGGATAACCGGTCGCTGTCAT 43


RESULT 11
AR365669
LOCUS       AR365669                  33 bp    DNA     linear   PAT 03-SEP-2003
DEFINITION  Sequence 36 from patent US 5519127.
ACCESSION   AR365669
VERSION     AR365669.1  GI:34429581
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 33)
  AUTHORS   Shah,J., Buharin,A. and Lane,D.J.
  TITLE     Nucleic acid probes for the detection of Pneumocystis carinii
  JOURNAL   Patent: US 5519127-A 36 21-MAY-1996;
FEATURES             Location/Qualifiers
     source          1. .33
                     /organism="unknown"
                     /mol_type="genomic DNA"
ORIGIN

  Query Match             34.5%;  Score 15.2;  DB 6;  Length 33;
  Best Local Similarity   77.3%;  Pred. No. 2.2e+05;
  Matches   17;  Conservative    1;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          8 CCGTTCCTTCTTAATAACCGGT 29
              :| ||||||||  || ||||||
Db          3 YCCTTCCTTCTGGATTACCGGT 24


RESULT 12
AX147162
LOCUS       AX147162                  37 bp    DNA     linear   PAT 08-JUN-2001
DEFINITION  Sequence 4 from Patent WO0136457.
ACCESSION   AX147162
VERSION     AX147162.1  GI:14346333
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            other sequences; artificial sequences.
REFERENCE   1
  AUTHORS   Murdin,A.D., Oomen,R.P., Wang,J. and Dunn,P.
  TITLE     Chlamydia antigens and corresponding dna fragments and uses thereof
  JOURNAL   Patent: WO 0136457-A 4 25-MAY-2001;
            Aventis Pasteur Limited (CA)
FEATURES             Location/Qualifiers
     source          1. .37
                     /organism="synthetic construct"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32630"
                     /note="3' PCR primer"
ORIGIN

  Query Match             34.5%;  Score 15.2;  DB 6;  Length 37;
  Best Local Similarity   71.4%;  Pred. No. 2.2e+05;
  Matches   20;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Qy          2 CGGGTCCCGTTCCTTCTTAATAACCGGT 29
              ||| ||||| || || |  |||| || |
Db          5 CGGATCCCGCTCTTTTTAGATAAGCGTT 32


RESULT 13
AX081593/c
LOCUS       AX081593                  47 bp    DNA     linear   PAT 27-FEB-2001
DEFINITION  Sequence 98 from Patent WO0109350.
ACCESSION   AX081593
VERSION     AX081593.1  GI:13170418
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            other sequences; artificial sequences.
REFERENCE   1
  AUTHORS   Berthet,F.X., Dalemans,W.L., Denoel,P., Dequesne,G.S., Feron,C.S.,
            Lobet,Y.S., Poolman,J.S., Thiry,G.S., Lhonnard,J.S. and Voet,P.S.
  TITLE     Genetically engineered bleb vaccine
  JOURNAL   Patent: WO 0109350-A 98 08-FEB-2001;
            SMITHKLINE BEECHAM BIOLOGICALS S.A. (BE)
FEATURES             Location/Qualifiers
     source          1. .47
                     /organism="synthetic construct"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32630"
                     /note="PNS4 primer"
```

```
ORIGIN

  Query Match           34.5%;  Score 15.2;  DB 6;  Length 47;
  Best Local Similarity  71.4%;  Pred. No. 2.2e+05;
  Matches   20;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Qy         17 CTTAATAACCGGTCGCGGTTATTAAGAA 44
              | || |  |||| |||| | |||||| |
Db         41 CATATTTTCCGGACGCGTTAATTAAGGA 14


RESULT 14
AX374774/c
LOCUS       AX374774                 47 bp    DNA     linear   PAT 01-MAR-2002
DEFINITION  Sequence 98 from Patent WO0209746.
ACCESSION   AX374774
VERSION     AX374774.1  GI:19169676
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            other sequences; artificial sequences.
REFERENCE   1
  AUTHORS   Berthet,J.G., Dalemans,W.G., Denoel,P.G., Dequesne,G.G.,
            Feron,C.G., Garcon,N.G., Lobet,Y.G., Poolman,J.G., Thiry,G.G.,
            Thonnard,J.G. and Voet,P.G.
  TITLE     Vaccine composition
  JOURNAL   Patent: WO 0209746-A 98 07-FEB-2002;
            SMITHKLINE BEECHAM BIOLOGICALS S.A. (BE)
FEATURES             Location/Qualifiers
     source          1. .47
                     /organism="synthetic construct"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32630"
                     /note="PNS4 primer"
ORIGIN

  Query Match           34.5%;  Score 15.2;  DB 6;  Length 47;
  Best Local Similarity  71.4%;  Pred. No. 2.2e+05;
  Matches   20;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Qy         17 CTTAATAACCGGTCGCGGTTATTAAGAA 44
              | || |  |||| |||| | |||||| |
Db         41 CATATTTTCCGGACGCGTTAATTAAGGA 14


RESULT 15
AX248444/c
LOCUS       AX248444                 31 bp    DNA     linear   PAT 28-SEP-2001
DEFINITION  Sequence 523 from Patent WO0166800.
ACCESSION   AX248444
VERSION     AX248444.1  GI:15863067
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Cargill,M., Ireland,J.S. and Lander,E.S.
  TITLE     Human single nucleotide polymorphisms
  JOURNAL   Patent: WO 0166800-A 523 13-SEP-2001;
            WHITEHEAD INSTITUTE FOR BIOMEDICAL RESEARCH (US)
FEATURES             Location/Qualifiers
     source          1. .31
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

  Query Match           34.1%;  Score 15;  DB 6;  Length 31;
  Best Local Similarity  78.3%;  Pred. No. 2.7e+05;
  Matches   18;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy         12 TCCTTCTTAATAACCGGTCGCGG 34
                || |||||||| | | |||| ||
Db         29 TCTTTCTTAATGAKCTGTCGGGG 7


RESULT 16
MMU41925
LOCUS       MMU41925                 50 bp    DNA     linear   ROD 05-JAN-1996
DEFINITION  Mus musculus recombination between immunoglobulin heavy chain and
            c-myc.
ACCESSION   U41925
VERSION     U41925.1  GI:1147659
KEYWORDS    .
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 50)
  AUTHORS   Muller,J.R.
  TITLE     Direct Submission
  JOURNAL   Submitted (05-DEC-1995) Jurgen R. Muller, Lab of Genetics, NIH/NCI,
            Bldg. 37, Room 2B09, 37 Convent Dr., Bethesda, MD 20892-4255, USA
FEATURES             Location/Qualifiers
     source          1. .50
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /strain="BALB/cAn"
                     /db_xref="taxon:10090"
                     /chromosome="T(12;15)"
                     /map="T(12f1;15d2)"
                     /tissue_type="oil granuloma 7 days post pristane"
                     /dev_stage="7 days post pristane"
ORIGIN

  Query Match           34.1%;  Score 15;  DB 10;  Length 50;
  Best Local Similarity  67.7%;  Pred. No. 2.7e+05;
  Matches   21;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy          4 GGTCCCGTTCCTTCTTAATAACCGGTCGCGG 34
              ||  ||  ||||||| | | ||| ||| | |
Db         11 GGGACCAGTCCTTCTGACTTACCAGTCTCTG 41


RESULT 17
AX582316/c
LOCUS       AX582316                 31 bp    DNA     linear   PAT 10-JAN-2003
DEFINITION  Sequence 4154 from Patent WO0211674.
ACCESSION   AX582316
VERSION     AX582316.1  GI:27654126
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            other sequences; artificial sequences.
REFERENCE   1
  AUTHORS   Thompson,J., Mcswiggen,J., Mckenzie,T., Ayers,D., Szymkowski,D.E.
            and Grupe,A.
  TITLE     Method and reagent for the inhibition of calcium activated chloride
            channel-1 (clca-1)
  JOURNAL   Patent: WO 0211674-A 4154 14-FEB-2002;
            RIBOZYME PHARMACEUTICALS, INC. (US) ; Syntex (U.S.A.) LLC (US) ;
            Thompson, James (US)
FEATURES             Location/Qualifiers
     source          1. .31
                     /organism="synthetic construct"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32630"
                     /note="Enzymatic Nucleic Acid"
ORIGIN

  Query Match           33.6%;  Score 14.8;  DB 6;  Length 31;
  Best Local Similarity  73.1%;  Pred. No. 3.3e+05;
  Matches   19;  Conservative    0;  Mismatches    7;  Indels    0;  Gaps    0;

Qy          5 GTCCCGTTCCTTCTTAATAACCGGTC 30
```

```
              ||||||||| || |   || || |||
Db         30 GTCCCGTTCGTTGTAGCTAGCCCGTC 5


RESULT 18
I03806/c
LOCUS       I03806                    33 bp    DNA     linear   PAT 02-DEC-1994
DEFINITION  Sequence 9 from Patent EP 0055942.
ACCESSION   I03806
VERSION     I03806.1  GI:592012
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 33)
  AUTHORS   Inouye,M. and Nakamura,K.
  TITLE     Plasmid cloning vehicles
  JOURNAL   Patent: EP 0055942-A2 9 14-JUL-1982;
FEATURES             Location/Qualifiers
     source          1. .33
                     /organism="unknown"
                     /mol_type="unassigned DNA"
ORIGIN

  Query Match           33.6%;  Score 14.8;  DB 6;  Length 33;
  Best Local Similarity   73.1%;  Pred. No. 3.3e+05;
  Matches   19;  Conservative    0;  Mismatches    7;  Indels    0;  Gaps    0;

Qy          6 TCCCGTTCCTTCTTAATAACCGGTCG 31
              |||| ||| || |||||| ||  | |
Db         29 TCCCTTTCATTATTAATACCCTCTAG 4


RESULT 19
AR429930
LOCUS       AR429930                  36 bp    DNA     linear   PAT 18-DEC-2003
DEFINITION  Sequence 40 from patent US 6645765.
ACCESSION   AR429930
VERSION     AR429930.1  GI:40190357
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 36)
  AUTHORS   Anderson,H.M., Chay,C.A., Chen,G. and Conner,T.W.
  TITLE     Plant regulatory sequences for control of gene expression
  JOURNAL   Patent: US 6645765-A 40 11-NOV-2003;
FEATURES             Location/Qualifiers
     source          1. .36
                     /organism="unknown"
                     /mol_type="genomic DNA"
ORIGIN

  Query Match           33.6%;  Score 14.8;  DB 6;  Length 36;
  Best Local Similarity   73.1%;  Pred. No. 3.3e+05;
  Matches   19;  Conservative    0;  Mismatches    7;  Indels    0;  Gaps    0;

Qy          9 CGTTCCTTCTTAATAACCGGTCGCGG 34
              | || ||||| | | | |||| ||||
Db          6 CCTTTCTTCTCACTCAGCGGTTGCGG 31


RESULT 20
AX099684
LOCUS       AX099684                  36 bp    DNA     linear   PAT 02-APR-2001
DEFINITION  Sequence 40 from Patent WO0119976.
ACCESSION   AX099684
VERSION     AX099684.1  GI:13538738
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            other sequences; artificial sequences.
REFERENCE   1
  AUTHORS   Anderson,H.M., Chay,C.A., Chen,G. and Conner,T.W.
  TITLE     Plant regulatory sequences for control of gene expression
  JOURNAL   Patent: WO 0119976-A 40 22-MAR-2001;
            MONSANTO COMPANY (US)
FEATURES             Location/Qualifiers
     source          1. .36
                     /organism="synthetic construct"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32630"
                     /note="primer"
ORIGIN

  Query Match           33.6%;  Score 14.8;  DB 6;  Length 36;
  Best Local Similarity   73.1%;  Pred. No. 3.3e+05;
  Matches   19;  Conservative    0;  Mismatches    7;  Indels    0;  Gaps    0;

Qy          9 CGTTCCTTCTTAATAACCGGTCGCGG 34
              | || ||||| | | | |||| ||||
Db          6 CCTTTCTTCTCACTCAGCGGTTGCGG 31


RESULT 21
I03802/c
LOCUS       I03802                    37 bp    DNA     linear   PAT 02-DEC-1994
DEFINITION  Sequence 5 from Patent EP 0055942.
ACCESSION   I03802
VERSION     I03802.1  GI:592010
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 37)
  AUTHORS   Inouye,M. and Nakamura,K.
  TITLE     Plasmid cloning vehicles
  JOURNAL   Patent: EP 0055942-A2 5 14-JUL-1982;
FEATURES             Location/Qualifiers
     source          1. .37
                     /organism="unknown"
                     /mol_type="unassigned DNA"
ORIGIN

  Query Match           33.6%;  Score 14.8;  DB 6;  Length 37;
  Best Local Similarity   73.1%;  Pred. No. 3.3e+05;
  Matches   19;  Conservative    0;  Mismatches    7;  Indels    0;  Gaps    0;

Qy          6 TCCCGTTCCTTCTTAATAACCGGTCG 31
              |||| ||| || |||||| ||  | |
Db         29 TCCCTTTCATTATTAATACCCTCTAG 4


RESULT 22
AR288505/c
LOCUS       AR288505                  47 bp    DNA     linear   PAT 12-JUN-2003
DEFINITION  Sequence 240 from patent US 6537751.
ACCESSION   AR288505
VERSION     AR288505.1  GI:31675789
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 47)
  AUTHORS   Cohen,D., Chumakov,I. and Blumenfeld,M.
  TITLE     Biallelic markers for use in constructing a high density
            disequilibrium map of the human genome
  JOURNAL   Patent: US 6537751-A 240 25-MAR-2003;
FEATURES             Location/Qualifiers
     source          1. .47
                     /organism="unknown"
                     /mol_type="genomic DNA"
ORIGIN

  Query Match           33.6%;  Score 14.8;  DB 6;  Length 47;
  Best Local Similarity   61.1%;  Pred. No. 3.3e+05;
```

```
  Matches   22;  Conservative    1;  Mismatches   13;  Indels    0;  Gaps  0;

Qy          9 CGTTCCTTCTTAATAACCGGTCGCGGTTATTAAGAA 44
              | ||   | ||||||   | ||:| ||| | || ||
Db         46 CATTTAATTTTAATACATGCTCRCTGTTTTGAAAAA 11


RESULT 23
AR289990/c
LOCUS       AR289990                  47 bp    DNA     linear   PAT 12-JUN-2003
DEFINITION  Sequence 1725 from patent US 6537751.
ACCESSION   AR289990
VERSION     AR289990.1  GI:31677274
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 47)
  AUTHORS   Cohen,D., Chumakov,I. and Blumenfeld,M.
  TITLE     Biallelic markers for use in constructing a high density
            disequilibrium map of the human genome
  JOURNAL   Patent: US 6537751-A 1725 25-MAR-2003;
FEATURES             Location/Qualifiers
     source          1. .47
                     /organism="unknown"
                     /mol_type="genomic DNA"
ORIGIN

  Query Match             33.6%;  Score 14.8;  DB 6;  Length 47;
  Best Local Similarity   59.5%;  Pred. No. 3.3e+05;
  Matches   25;  Conservative    0;  Mismatches   17;  Indels    0;  Gaps    0;

Qy          3 GGGTCCCGTTCCTTCTTAATAACCGGTCGCGGTTATTAAGAA 44
              || ||||  ||  |||| ||||     | |  ||| | | ||
Db         46 GGCTCCCCATCTCTCTTTATAARTTAGCACCATTAATTATAA 5


RESULT 24
AR032973/c
LOCUS       AR032973                  50 bp    DNA     linear   PAT 29-SEP-1999
DEFINITION  Sequence 585 from patent US 5869241.
ACCESSION   AR032973
VERSION     AR032973.1  GI:5948578
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 50)
  AUTHORS   Edwards,C.A., Cantor,C.R., Andrews,B.M., Turin,L.M. and Fry,K.E.
  TITLE     Method of determining DNA sequence preference of a DNA-binding
            molecule
  JOURNAL   Patent: US 5869241-A 585 09-FEB-1999;
FEATURES             Location/Qualifiers
     source          1. .50
                     /organism="unknown"
                     /mol_type="unassigned DNA"
ORIGIN

  Query Match             33.6%;  Score 14.8;  DB 6;  Length 50;
  Best Local Similarity   73.1%;  Pred. No. 3.3e+05;
  Matches   19;  Conservative    0;  Mismatches    7;  Indels    0;  Gaps    0;

Qy         11 TTCCTTCTTAATAACCGGTCGCGGTT 36
              | | ||  || ||||||||  |||||
Db         27 TGCTTTTATACTAACCGGTTTCGGTT 2


RESULT 25
I29713/c
LOCUS       I29713                    50 bp    DNA     linear   PAT 06-FEB-1997
DEFINITION  Sequence 585 from patent US 5578444.
ACCESSION   I29713
VERSION     I29713.1  GI:1820504
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 50)
  AUTHORS   Edwards,C.A., Cantor,C.R., Andrews,B.M., Turin,L.M. and Fry,K.E.
  TITLE     Sequence-directed DNA-binding molecules compositions and methods
  JOURNAL   Patent: US 5578444-A 585 26-NOV-1996;
FEATURES             Location/Qualifiers
     source          1. .50
                     /organism="unknown"
                     /mol_type="unassigned DNA"
ORIGIN

  Query Match             33.6%;  Score 14.8;  DB 6;  Length 50;
  Best Local Similarity   73.1%;  Pred. No. 3.3e+05;
  Matches   19;  Conservative    0;  Mismatches    7;  Indels    0;  Gaps    0;

Qy         11 TTCCTTCTTAATAACCGGTCGCGGTT 36
              | | ||  || ||||||||  |||||
Db         27 TGCTTTTATACTAACCGGTTTCGGTT 2


RESULT 26
I91387/c
LOCUS       I91387                    50 bp    DNA     linear   PAT 01-DEC-1998
DEFINITION  Sequence 585 from patent US 5726014.
ACCESSION   I91387
VERSION     I91387.1  GI:3935857
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 50)
  AUTHORS   Edwards,C.A., Cantor,C.R., Andrews,B.M. and Turin,L.M.
  TITLE     Screening assay for the detection of DNA-binding molecules
  JOURNAL   Patent: US 5726014-A 585 10-MAR-1998;
FEATURES             Location/Qualifiers
     source          1. .50
                     /organism="unknown"
                     /mol_type="unassigned DNA"
ORIGIN

  Query Match             33.6%;  Score 14.8;  DB 6;  Length 50;
  Best Local Similarity   73.1%;  Pred. No. 3.3e+05;
  Matches   19;  Conservative    0;  Mismatches    7;  Indels    0;  Gaps    0;

Qy         11 TTCCTTCTTAATAACCGGTCGCGGTT 36
              | | ||  || ||||||||  |||||
Db         27 TGCTTTTATACTAACCGGTTTCGGTT 2


RESULT 27
AR209637/c
LOCUS       AR209637                  50 bp    DNA     linear   PAT 20-JUN-2002
DEFINITION  Sequence 585 from patent US 6384208.
ACCESSION   AR209637
VERSION     AR209637.1  GI:21511114
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 50)
  AUTHORS   Edwards,C.A., Cantor,C.R., Andrews,B.M., Turin,L.M. and Fry,K.E.
  TITLE     Sequence directed DNA binding molecules compositions and methods
  JOURNAL   Patent: US 6384208-A 585 07-MAY-2002;
FEATURES             Location/Qualifiers
     source          1. .50
                     /organism="unknown"
                     /mol_type="unassigned DNA"
ORIGIN

  Query Match             33.6%;  Score 14.8;  DB 6;  Length 50;
```

```
  Best Local Similarity   73.1%;  Pred. No. 3.3e+05;
  Matches   19;  Conservative    0;  Mismatches    7;  Indels    0;  Gaps    0;

Qy         11 TTCCTTCTTAATAACCGGTCGCGGTT 36
              | | ||  || ||||||||  |||||
Db         27 TGCTTTTATACTAACCGGTTTCGGTT 2


RESULT 28
CQ868809
LOCUS       CQ868809                  32 bp    DNA     linear   PAT 13-SEP-2004
DEFINITION  Sequence 20 from Patent WO2004073728.
ACCESSION   CQ868809
VERSION     CQ868809.1  GI:51998743
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            other sequences; artificial sequences.
REFERENCE   1
  AUTHORS   Westphal,O., Waelli,T., Gorczynski,R., Mueller,S., Mach,J.P.,
            Hartmann,A., Bessler,W., Hofmann,P., Zaehringer,U., Alexander,C.,
            vor dem Esche,U., Ulmer,A. and Verdini,A.
  TITLE     Compositions comprising fetal hemoglobin and bacterial endotoxin
            and optionally additional fetal liver components
  JOURNAL   Patent: WO 2004073728-A 20 02-SEP-2004;
            Clinique La Prairie Research SA (LU)
FEATURES             Location/Qualifiers
     source          1. .32
                     /organism="synthetic construct"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32630"
ORIGIN

  Query Match            33.2%;  Score 14.6;  DB 6;  Length 32;
  Best Local Similarity   81.0%;  Pred. No. 4.1e+05;
  Matches   17;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy         19 TAATAACCGGTCGCGGTTATT 39
              ||||||||||   ||| |||
Db          1 TAATAACCGGTATGGGTCATT 21


RESULT 29
AR365650
LOCUS       AR365650                  33 bp    DNA     linear   PAT 03-SEP-2003
DEFINITION  Sequence 17 from patent US 5519127.
ACCESSION   AR365650
VERSION     AR365650.1  GI:34429562
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 33)
  AUTHORS   Shah,J., Buharin,A. and Lane,D.J.
  TITLE     Nucleic acid probes for the detection of Pneumocystis carinii
  JOURNAL   Patent: US 5519127-A 17 21-MAY-1996;
FEATURES             Location/Qualifiers
     source          1. .33
                     /organism="unknown"
                     /mol_type="genomic DNA"
ORIGIN

  Query Match            33.2%;  Score 14.6;  DB 6;  Length 33;
  Best Local Similarity   81.0%;  Pred. No. 4.1e+05;
  Matches   17;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          9 CGTTCCTTCTTAATAACCGGT 29
              | ||||||||  || ||||||
Db          4 CCTTCCTTCTGGATTACCGGT 24


RESULT 30
AR365637/c
LOCUS       AR365637                  34 bp    DNA     linear   PAT 03-SEP-2003
DEFINITION  Sequence 4 from patent US 5519127.
ACCESSION   AR365637
VERSION     AR365637.1  GI:34429549
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 34)
  AUTHORS   Shah,J., Buharin,A. and Lane,D.J.
  TITLE     Nucleic acid probes for the detection of Pneumocystis carinii
  JOURNAL   Patent: US 5519127-A 4 21-MAY-1996;
FEATURES             Location/Qualifiers
     source          1. .34
                     /organism="unknown"
                     /mol_type="genomic DNA"
ORIGIN

  Query Match            33.2%;  Score 14.6;  DB 6;  Length 34;
  Best Local Similarity   81.0%;  Pred. No. 4.1e+05;
  Matches   17;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          9 CGTTCCTTCTTAATAACCGGT 29
              | ||||||||  || ||||||
Db         31 CCTTCCTTCTGGATTACCGGT 11


RESULT 31
AX739899/c
LOCUS       AX739899                  35 bp    DNA     linear   PAT 08-MAY-2003
DEFINITION  Sequence 7 from Patent WO03025016.
ACCESSION   AX739899
VERSION     AX739899.1  GI:30519182
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            other sequences; artificial sequences.
REFERENCE   1
  AUTHORS   Dawbarn,D., Allen,S.J. and Robertson,A.G.
  TITLE     Polypeptide purification method
  JOURNAL   Patent: WO 03025016-A 7 27-MAR-2003;
            The University of Bristol (GB)
FEATURES             Location/Qualifiers
     source          1. .35
                     /organism="synthetic construct"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32630"
ORIGIN

  Query Match            33.2%;  Score 14.6;  DB 6;  Length 35;
  Best Local Similarity   73.9%;  Pred. No. 4.1e+05;
  Matches   17;  Conservative    1;  Mismatches    5;  Indels    0;  Gaps    0;

Qy         21 ATAACCGGTCGCGGTTATTAAGA 43
              | |||||||||  | : ||||| |
Db         32 AAAACCGGTCGGGYYCATTAATA 10


RESULT 32
AR365666
LOCUS       AR365666                  39 bp    DNA     linear   PAT 03-SEP-2003
DEFINITION  Sequence 33 from patent US 5519127.
ACCESSION   AR365666
VERSION     AR365666.1  GI:34429578
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 39)
  AUTHORS   Shah,J., Buharin,A. and Lane,D.J.
  TITLE     Nucleic acid probes for the detection of Pneumocystis carinii
  JOURNAL   Patent: US 5519127-A 33 21-MAY-1996;
FEATURES             Location/Qualifiers
```

```
     source          1. .39
                     /organism="unknown"
                     /mol_type="genomic DNA"
ORIGIN

  Query Match            33.2%;  Score 14.6;  DB 6;  Length 39;
  Best Local Similarity  81.0%;  Pred. No. 4e+05;
  Matches   17;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          9 CGTTCCTTCTTAATAACCGGT 29
              | ||||||||  || ||||||
Db          5 CCTTCCTTCTGGATTACCGGT 25


RESULT 33
AX642259/c
LOCUS       AX642259                 45 bp    DNA     linear   PAT 21-FEB-2003
DEFINITION  Sequence 16 from Patent WO02055697.
ACCESSION   AX642259
VERSION     AX642259.1  GI:28474706
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            other sequences; artificial sequences.
REFERENCE   1
  AUTHORS   Kamb,C.A., Poritz,M.A. and Teng,D.H.
  TITLE     Human rhinovirus assays, and compositions comprising
            anti-rhinoviral perturbagens
  JOURNAL   Patent: WO 02055697-A 16 18-JUL-2002;
            Deltagen Proteomics, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .45
                     /organism="synthetic construct"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32630"
                     /note="rh1D.R2 primer"
ORIGIN

  Query Match            33.2%;  Score 14.6;  DB 6;  Length 45;
  Best Local Similarity  62.2%;  Pred. No. 4e+05;
  Matches   23;  Conservative    0;  Mismatches   14;  Indels    0;  Gaps    0;

Qy          4 GGTCCCGTTCCTTCTTAATAACCGGTCGCGGTTATTA 40
              |||  | ||   || || ||| ||| ||| ||  | |
Db         37 GGTGACATTAAATCCTATTAAGCGGCCGCTGTATTGA 1


RESULT 34
I13724
LOCUS       I13724                   46 bp    DNA     linear   PAT 26-SEP-1995
DEFINITION  Sequence 30 from patent US 5439829.
ACCESSION   I13724
VERSION     I13724.1  GI:996790
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 46)
  AUTHORS   Anderson,L.D., Cook,J.A., David,G.S., Hochschwender,S.M.,
            Kasher,M.S., Smith,M.C. and Stemmer,W.P.C.
  TITLE     Immobilization of biologically active molecules by changing the
            Oxidation state of a chelated transition metal ion
  JOURNAL   Patent: US 5439829-A 30 08-AUG-1995;
FEATURES             Location/Qualifiers
     source          1. .46
                     /organism="unknown"
                     /mol_type="unassigned DNA"
ORIGIN

  Query Match            33.2%;  Score 14.6;  DB 6;  Length 46;
  Best Local Similarity  69.0%;  Pred. No. 4e+05;
  Matches   20;  Conservative    0;  Mismatches    9;  Indels    0;  Gaps    0;

Qy         16 TCTTAATAACCGGTCGCGGTTATTAAGAA 44
              | |||||||    ||||  |||| ||| |
Db          9 TATTAATAATGTATCGCTATTATAAAGGA 37


RESULT 35
AR291680
LOCUS       AR291680                 47 bp    DNA     linear   PAT 12-JUN-2003
DEFINITION  Sequence 3415 from patent US 6537751.
ACCESSION   AR291680
VERSION     AR291680.1  GI:31678964
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 47)
  AUTHORS   Cohen,D., Chumakov,I. and Blumenfeld,M.
  TITLE     Biallelic markers for use in constructing a high density
            disequilibrium map of the human genome
  JOURNAL   Patent: US 6537751-A 3415 25-MAR-2003;
FEATURES             Location/Qualifiers
     source          1. .47
                     /organism="unknown"
                     /mol_type="genomic DNA"
ORIGIN

  Query Match            33.2%;  Score 14.6;  DB 6;  Length 47;
  Best Local Similarity  64.5%;  Pred. No. 4e+05;
  Matches   20;  Conservative    1;  Mismatches   10;  Indels    0;  Gaps    0;

Qy         10 GTTCCTTCTTAATAACCGGTCGCGGTTATTA 40
              |||| | |:||| ||||  || | | |  ||
Db         16 GTTCATCCKTAAAAACCATTCTCAGCTCCTA 46


RESULT 36
AX462584
LOCUS       AX462584                 20 bp    DNA     linear   PAT 15-JUL-2002
DEFINITION  Sequence 328 from Patent EP1217079.
ACCESSION   AX462584
VERSION     AX462584.1  GI:21885797
KEYWORDS    .
SOURCE      Aegilops tauschii
  ORGANISM  Aegilops tauschii
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Pooideae; Triticeae; Aegilops.
REFERENCE   1
  AUTHORS   Bernard,M., Sourdille,P. and Guyomarch,H.
  TITLE     Microsatellite markers from Triticum tauschii
  JOURNAL   Patent: EP 1217079-A 328 26-JUN-2002;
            INSTITUT NATIONAL DE LA RECHERCHE AGRONOMIQUE (INRA) (FR)
FEATURES             Location/Qualifiers
     source          1. .20
                     /organism="Aegilops tauschii"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:37682"
ORIGIN

  Query Match            32.7%;  Score 14.4;  DB 6;  Length 20;
  Best Local Similarity  93.8%;  Pred. No. 5e+05;
  Matches   15;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy         13 CCTTCTTAATAACCGG 28
              ||||||||||  |||||
Db          1 CCTTCTTAATGACCGG 16


RESULT 37
BD061913
LOCUS       BD061913                 32 bp    DNA     linear   PAT 27-AUG-2002
DEFINITION  Antigenic composition and method of detection for Helicobacter
            Pylori.
```

```
ACCESSION   BD061913
VERSION     BD061913.1  GI:22607518
KEYWORDS    JP 2001517091-A/247.
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            other sequences; artificial sequences.
REFERENCE   1  (bases 1 to 32)
  AUTHORS   Chow,T.P., Fry,K.E., Lim,M.Y. and Mcatee,C.P.
  TITLE     Antigenic composition and method of detection for Helicobacter
  JOURNAL   Patent: JP 2001517091-A 247 02-OCT-2001;
            GENELABS TECHNOLOGIES INC
COMMENT     PN   JP 2001517091-A/247
            PD   02-OCT-2001
            PF   25-APR-1998 JP 1998547263
            PR   25-APR-1997 US    60/045107,14-OCT-1997 US    60/061958 PI
            THERESA P CHOW,KIRK E FRY,MOON Y LIM,C P MCATEE PC
            C12N15/31,C07K14/205,C07K16/12,A61K39/106
            CC   Strandedness: Single;
            CC   Topology: Linear;
            FH   Key             Location/Qualifiers.
FEATURES              Location/Qualifiers
     source           1. .32
                      /organism="synthetic construct"
                      /mol_type="genomic DNA"
                      /db_xref="taxon:32630"
ORIGIN

  Query Match             32.7%;  Score 14.4;  DB 6;  Length 32;
  Best Local Similarity   75.0%;  Pred. No. 5e+05;
  Matches   18;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          1 GCGGGTCCCGTTCCTTCTTAATAA 24
              |||| |||| |||| |||||||
Db          4 GCGGATCCCATTCCACTATAATAA 27


RESULT 38
AX574343
LOCUS       AX574343                 34 bp    DNA     linear   PAT 07-JAN-2003
DEFINITION  Sequence 7 from Patent WO0250545.
ACCESSION   AX574343
VERSION     AX574343.1  GI:27551693
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            other sequences; artificial sequences.
REFERENCE   1
  AUTHORS   Pelletier,J., Gros,P. and Dubow,M.
  TITLE     Compositions and methods involving staphylococcus aureus protein
            staau-r9
  JOURNAL   Patent: WO 0250545-A 7 27-JUN-2002;
            Phagetech Inc. (CA)
FEATURES              Location/Qualifiers
     source           1. .34
                      /organism="synthetic construct"
                      /mol_type="unassigned DNA"
                      /db_xref="taxon:32630"
                      /note="Synthetic oligonucleotide"
ORIGIN

  Query Match             32.7%;  Score 14.4;  DB 6;  Length 34;
  Best Local Similarity   75.0%;  Pred. No. 5e+05;
  Matches   18;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          1 GCGGGTCCCGTTCCTTCTTAATAA 24
              |||| ||||  ||||| || || |
Db          2 GCGGATCCCTATCCTTTTTCATTA 25


RESULT 39
BD235752
LOCUS       BD235752                 40 bp    DNA     linear   PAT 17-JUL-2003
DEFINITION  Strengthened functional expression of heterogenous G
            protein-coupled receptor.
ACCESSION   BD235752
VERSION     BD235752.1  GI:33045522
KEYWORDS    JP 2002523090-A/5.
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            other sequences; artificial sequences.
REFERENCE   1  (bases 1 to 40)
  AUTHORS   Pausch,M.H., Lai,M., Silverman,S., Birsan,C., Baumbauch,W.,
            Tseng,E., Kajkowski,E.M. and Ozenberger,B.A.
  TITLE     Strengthened functional expression of heterogenous G
            protein-coupled receptor
  JOURNAL   Patent: JP 2002523090-A 5 30-JUL-2002;
            BASF AG
COMMENT     OS   Artificial Sequence
            PN   JP 2002523090-A/5
            PD   30-JUL-2002
            PF   01-SEP-1999 JP 2000567691
            PR   01-SEP-1998 US    60/098704
            PI   MARK HENRY PAUSCH,MARGARET LAI,SANFORD SILVERMAN,CAMELIA  PI
            BIRSAN,
            PI   WILLIAM BAUMBAUCH,EUGENE TSENG,EILEEN MARIE KAJKOWSKI, PI
            BRADLEY ALTON OZENBERGER
            PC   C12N1/19,C07K14/72,C12N15/09,C12P21/02,C12Q1/02,G01N33/15, PC
             G01N33/50,
            PC   G01N33/566//(C12N1/19,C12R1:865),(C12P21/02,C12R1:865),C12N15/
            PC   00
            CC   Description of Artificial Sequence:Oligonucleotide FH   Key
                 Location/Qualifiers
            FT   source          1. .40
            FT                   /organism='Artificial Sequence'.
FEATURES              Location/Qualifiers
     source           1. .40
                      /organism="synthetic construct"
                      /mol_type="genomic DNA"
                      /db_xref="taxon:32630"
ORIGIN

  Query Match             32.7%;  Score 14.4;  DB 6;  Length 40;
  Best Local Similarity   65.6%;  Pred. No. 5e+05;
  Matches   21;  Conservative    0;  Mismatches   11;  Indels    0;  Gaps    0;

Qy          9 CGTTCCTTCTTAATAACCGGTCGCGGTTATTA 40
              || | | | ||| | |||||| | | || |
Db          5 CGCTGCCTTTTACTTACCGGTCACCATCATGA 36


RESULT 40
BD235803
LOCUS       BD235803                 40 bp    DNA     linear   PAT 17-JUL-2003
DEFINITION  Method of modifying function of heterogenous G protein-coupled
            receptor.
ACCESSION   BD235803
VERSION     BD235803.1  GI:33045573
KEYWORDS    JP 2002523091-A/5.
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            other sequences; artificial sequences.
REFERENCE   1  (bases 1 to 40)
  AUTHORS   Pausch,M.H. and Wess,J.
  TITLE     Method of modifying function of heterogenous G protein-coupled
  JOURNAL   Patent: JP 2002523091-A 5 30-JUL-2002;
            BASF AG
COMMENT     OS   Artificial Sequence
            PN   JP 2002523091-A/5
            PD   30-JUL-2002
            PF   01-SEP-1999 JP 2000567692
            PR   01-SEP-1998 US    60/098704
            PI   MARK HENRY PAUSCH,JURGEN WESS
            PC   C12N15/09,C07K14/72,C12N1/19,C12Q1/02,G01N33/15,G01N33/50, PC
             G01N33/566,
            PC   G01N33/68//(C12N1/19,C12R1:865),C12N15/00
            CC   Description of Artificial Sequence:oligonucleotide FH   Key
```

```
                     Location/Qualifiers
           FT   source          1. .40
           FT                   /organism='Artificial Sequence'.
FEATURES             Location/Qualifiers
     source          1. .40
                     /organism="synthetic construct"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:32630"
ORIGIN

  Query Match             32.7%;  Score 14.4;  DB 6;  Length 40;
  Best Local Similarity   65.6%;  Pred. No. 5e+05;
  Matches   21;  Conservative    0;  Mismatches   11;  Indels    0;  Gaps    0;

Qy          9 CGTTCCTTCTTAATAACCGGTCGCGGTTATTA 40
              || | | | ||| | ||||||| |  | || |
Db          5 CGCTGCCTTTTACTTACCGGTCACCATCATGA 36


Search completed: May 24, 2005, 12:43:56
Job time : 1770 secs
```

This Page Blank (usp

OM nucleic - nucleic search, using sw model

Run on: May 24, 2005, 08:29:04 ; Search time 258 Seconds
(without alignments)
1009.568 Million cell updates/sec

Title: US-10-673-063-3_COPY_900_943
Perfect score: 44
Sequence: 1 gcgggtcccgttccttctta.....ccggtcgcggttattaagaa 44

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 4167226

Minimum DB seq length: 0
Maximum DB seq length: 50

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 100 summaries

Database : N_Geneseq_16Dec04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004as:*
13: geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
c	1	17.4	39.5	39	13	ADS87007	Ads87007 PCR prime
c	2	16.2	36.8	32	6	AAL43407	Aal43407 SIMK muta
c	3	16	36.4	27	6	AAS20859	Aas20859 Gene-spec
	4	16	36.4	27	6	AAS20860	Aas20860 Gene-spec
	5	16	36.4	30	12	ADE85746	Ade85746 EphA2 inv
c	6	16	36.4	33	12	ADL27336	Adl27336 Forward p
c	7	15.8	35.9	45	2	AAQ87806	Aaq87806 Primer us
c	8	15.6	35.5	33	6	ABK89998	Abk89998 Human hea
	9	15.4	35.0	34	2	AAX27543	Aax27543 Staphylok
	10	15.2	34.5	37	5	AAF83844	Aaf83844 C. pneumo
c	11	15.2	34.5	38	9	ADB81046	Adb81046 LINE retr
c	12	15.2	34.5	47	3	AAZ65893	Aaz65893 Human map
c	13	15.2	34.5	47	4	AAF91472	Aaf91472 N. mening
c	14	15.2	34.5	47	6	ABK37852	Abk37852 pCMK(+)' c
c	15	15	34.1	31	4	AAI30035	Aai30035 Human sin
	16	15	34.1	47	3	AAZ69059	Aaz69059 Human map
	17	15	34.1	50	6	ABZ02614	Abz02614 Human leu
c	18	15	34.1	50	6	ABZ07725	Abz07725 Human leu
c	19	14.8	33.6	31	6	ABK59783	Abk59783 Human CLC
	20	14.8	33.6	36	4	AAF81443	Aaf81443 PCR prime
c	21	14.8	33.6	39	6	AAD43960	Aad43960 DI228 oli
c	22	14.8	33.6	47	3	AAZ67378	Aaz67378 Human map
c	23	14.8	33.6	50	2	AAQ69835	Aaq69835 Human pap
c	24	14.8	33.6	50	2	AAT64297	Aat64297 HPV type
c	25	14.8	33.6	50	2	AAX17585	Aax17585 Test sequ
c	26	14.8	33.6	50	6	ABK83076	Abk83076 DNA bindi
	27	14.8	33.6	50	6	ABZ02818	Abz02818 Human leu
c	28	14.8	33.6	50	12	ADE80615	Ade80615 Duplex ol
c	29	14.8	33.6	50	13	ADS91869	Ads91869 Nematode
	30	14.6	33.2	25	9	ACI97514	Aci97514 Human mic
c	31	14.6	33.2	34	2	AAQ10823	Aaq10823 Pneumocys
c	32	14.6	33.2	34	2	AAT42447	Aat42447 Probe 148
c	33	14.6	33.2	35	10	ADI35720	Adi35720 Human tyr
c	34	14.6	33.2	38	2	AAV06323	Aav06323 Human Col
	35	14.6	33.2	41	6	ABZ47662	Abz47662 Human ATP
	36	14.6	33.2	41	6	ABZ45067	Abz45067 Human ATP
c	37	14.6	33.2	41	12	ADL60971	Adl60971 Human org
c	38	14.6	33.2	45	6	ABN84317	Abn84317 Rhinoviru
	39	14.6	33.2	47	12	ADO18016	Ado18016 Primer of
	40	14.6	33.2	47	12	ADO18088	Ado18088 Primer of
	41	14.4	32.7	20	6	ABQ93060	Abq93060 T. tausch
c	42	14.4	32.7	25	9	ACI88894	Aci88894 Human mic
	43	14.4	32.7	29	10	ADK66726	Adk66726 Glucokina
	44	14.4	32.7	29	10	ADL09096	Adl09096 Staphyloc
	45	14.4	32.7	32	2	AAV90788	Aav90788 Primer Y1
	46	14.4	32.7	34	6	ABK87091	Abk87091 S. aureus
	47	14.4	32.7	40	3	AAZ94303	Aaz94303 Rat M3 mu
	48	14.4	32.7	40	3	AAZ94354	Aaz94354 Rat musca
	49	14.4	32.7	41	12	ADL60971	Adl60971 Human org
	50	14.4	32.7	47	3	AAZ66643	Aaz66643 Human map
c	51	14.4	32.7	50	5	AAC84632	Aac84632 SIV gene
c	52	14.4	32.7	50	10	ADD31889	Add31889 Butterfly
c	53	14.4	32.7	50	10	ADD31920	Add31920 Human ret
c	54	14.2	32.3	20	2	AAX97393	Aax97393 Primer us
c	55	14.2	32.3	20	8	ABZ81133	Abz81133 Dual spec
c	56	14.2	32.3	25	6	ABA92866	Aba92866 Human eos
c	57	14.2	32.3	25	8	ABV75348	Abv75348 Human ECP
	58	14.2	32.3	31	6	ABZ22006	Abz22006 Helicobac
	59	14.2	32.3	31	6	ABZ22037	Abz22037 Helicobac
c	60	14.2	32.3	31	6	AAS18580	Aas18580 Human tra
c	61	14.2	32.3	32	12	ADL70702	Adl70702 Escherich
c	62	14.2	32.3	33	6	AAH77221	Aah77221 PCR prime
	63	14.2	32.3	33	13	ADS73809	Ads73809 E. coli n
	64	14.2	32.3	34	2	AAV33572	Aav33572 Leukocyte
	65	14.2	32.3	36	2	AAQ62183	Aaq62183 Tryptopha
c	66	14.2	32.3	36	2	AAQ81847	Aaq81847 PD1 promo
c	67	14.2	32.3	38	3	AAZ43342	Aaz43342 Murine Ty
c	68	14.2	32.3	38	3	AAA05327	Aaa05327 PCR prime
	69	14.2	32.3	39	10	ADF54788	Adf54788 Influenza
c	70	14.2	32.3	41	6	ABL40260	Abl40260 Human pro
	71	14.2	32.3	41	10	ADC53158	Adc53158 Thermus t
c	72	14.2	32.3	45	6	AAL46572	Aal46572 Human FAP
c	73	14.2	32.3	45	6	AAL46580	Aal46580 Human FAP
c	74	14.2	32.3	48	10	ACF04958	Acf04958 Hair papi
c	75	14.2	32.3	49	5	ABA10693	Aba10693 Tail adap
	76	14.2	32.3	50	4	AAH90534	Aah90534 Human clo
c	77	14.2	32.3	50	6	ABZ04787	Abz04787 Human leu
c	78	14.2	32.3	50	10	ADG33460	Adg33460 Human DNA
c	79	14.2	32.3	50	12	ADH74736	Adh74736 LIC cloni
c	80	14	31.8	24	2	AAQ24974	Aaq24974 PCR prime
	81	14	31.8	24	4	AAS09446	Aas09446 Sense PCR
	82	14	31.8	30	4	AAH84364	Aah84364 Human cel
	83	14	31.8	32	3	AAZ36904	Aaz36904 PCR prime
	84	14	31.8	35	3	AAZ36906	Aaz36906 PCR prime
c	85	14	31.8	38	2	AAV54694	Aav54694 TRX2 gene
	86	14	31.8	38	2	AAZ31019	Aaz31019 Upstream
	87	14	31.8	38	4	AAF54760	Aaf54760 PCR prime
	88	14	31.8	38	10	AAD62934	Aad62934 S. aureus
c	89	14	31.8	41	2	AAV85621	Aav85621 LRP5 exon
c	90	14	31.8	42	11	ADM79708	Adm79708 Group B S
	91	14	31.8	45	10	ACC78190	Acc78190 DNA seque
	92	14	31.8	47	2	AAQ69500	Aaq69500 Human nuc
	93	14	31.8	47	2	AAT63962	Aat63962 Human nuc

94	14	31.8	47	2	AAX17250	Aax17250 Test sequ
95	14	31.8	47	6	ABK82741	Abk82741 DNA bindi
96	14	31.8	47	12	ADE80280	Ade80280 Duplex ol
c 97	13.8	31.4	24	6	ABQ07869	Abq07869 Oligonucl
c 98	13.8	31.4	24	6	ABQ01975	Abq01975 Oligonucl
99	13.8	31.4	24	6	ABQ07910	Abq07910 Oligonucl
c 100	13.8	31.4	25	4	AAH38015	Aah38015 SNP speci

ALIGNMENTS

RESULT 1
ADS87007/c
ID ADS87007 standard; DNA; 39 BP.
XX
AC ADS87007;
XX
DT 18-NOV-2004 (first entry)
XX
DE PCR primer 1 used to amplify murine TRP-2 cDNA.
XX
KW vaccine; ubiquitin; Ub; T-cell target; melanoma; sarcoma;
KW Hodgkins lymphoma; non-Hodgkins; leukaemia; neuroblastoma; myeloma;
KW lung cancer; stomach; skin; thyroid; ovary; prostate; womb; pancreas;
KW colon; bladder; breast; oesophagus; kidney; brain; mouse; murine; ss;
KW PCR; primer; TRP-2.
XX
OS Mus sp.
XX
PN WO2004035085-A1.
XX
PD 29-APR-2004.
XX
PF 16-OCT-2003; 2003WO-JP013279.
XX
PR 17-OCT-2002; 2002JP-00302816.
XX
PA (KYUS-) KYUSHU TLO CO LTD.
XX
PI Himeno K, Furue M, Maehara Y;
XX
DR WPI; 2004-357144/33.
XX
PT Gene vaccine containing cancer antigen genes ligated to ubiquitin genes
PT or cytokine genes for prevention and treatment of cancer.
XX
PS Example 6; SEQ ID NO 23; 266pp; Japanese.
XX
CC The invention relates to a novel genetic vaccine containing the ubiquitin
CC gene together with a gene encoding an antigenic protein containing a T-
CC cell target sequence. The vaccine of the invention may be useful for
CC prevention and treatment of cancers including melanoma, sarcoma, lymphoma
CC (Hodgkins or non-Hodgkins), leukaemia, neuroblastoma, myeloma and cancer
CC of the lung, stomach, skin, thyroid, ovary, prostate, womb, pancreas,
CC colon, bladder, breast, oesophagus, kidney or brain. The current sequence
CC is that of a genetic vaccine/ubiquitin (Ub)-related PCR primer of the
CC invention.
XX
SQ Sequence 39 BP; 11 A; 10 C; 12 G; 6 T; 0 U; 0 Other;

Query Match 39.5%; Score 17.4; DB 13; Length 39;
Best Local Similarity 68.6%; Pred. No. 1.4e+03;
Matches 24; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

```
Qy          5 GTCCCGTTCCTTCTTAATAACCGGTCGCGGTTATT 39
              | ||| | ||||||||||   ||| |||    |||
Db         36 GGCCCATGCCTTCTTAATTGGCGGCCGCTTGGATT 2
```

RESULT 2
AAL43407/c
ID AAL43407 standard; DNA; 32 BP.
XX
AC AAL43407;
XX
DT 02-SEP-2002 (first entry)
XX
DE SIMK mutagenic PCR primer (SIMKLOF).
XX
KW Salt stress-induced mitogen activated-protein kinase kinase; SIMKK; ss;
KW stress tolerance; salt tolerance; plant cell; PCR; primer; SIMK; SIMKLOF.
XX
OS Unidentified.
OS Synthetic.
XX
PN WO200238745-A2.
XX
PD 16-MAY-2002.
XX
PF 06-NOV-2001; 2001WO-EP012800.
XX
PR 07-NOV-2000; 2000AT-00001880.
XX
PA (OSTP) OESTERR FORSCH SEIBERSDORF.
XX
PI Meskiene I, Hirt H, Jonak C;
XX
DR WPI; 2002-490077/52.
XX
PT Salt stress-induced mitogen-activated protein kinase, useful for
PT improving stress tolerance, especially salt tolerance in eukaryotic
PT cells, particularly in plant cells.
XX
PS Example; Page 17; 42pp; English.
XX
CC The invention comprises the amino acid and coding sequence of a salt
CC stress-induced mitogen-activated protein kinase kinase (SIMKK). The SIMKK
CC DNA and protein sequences of the invention are useful for improving
CC stress tolerance (i.e. salt tolerance) in eukaryotic cells - particularly
CC in plant cells. The present DNA sequence represents a stress-induced
CC mitogen-activated protein kinase (SIMK) mutagenic PCR primer
XX
SQ Sequence 32 BP; 8 A; 9 C; 5 G; 10 T; 0 U; 0 Other;

Query Match 36.8%; Score 16.2; DB 6; Length 32;
Best Local Similarity 85.7%; Pred. No. 4.4e+03;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
Qy         23 AACCGGTCGCGGTTATTAAGA 43
              |||  |||||||||||  ||||
Db         31 AACATGTCGCGGTTATGAAGA 11
```

RESULT 3
AAS20859/c
ID AAS20859 standard; DNA; 27 BP.
XX
AC AAS20859;
XX
DT 09-APR-2002 (first entry)
XX
DE Gene-specific PCR primer Z-ST-P18 for cloning Z. marina sulfotransferase.
XX
KW Plant; transgenic; marine eelgrass; zosteric acid biosynthesis;
KW saline-resistance; anoxia-resistance; anti-fouling genetic trait;
KW marine vascular plant; sulphated phenolic compound; Zostera marina;
KW sulfotransferase; ST; enzyme; PCR; primer; ss.
XX
OS Zostera marina.
XX
PN WO200185971-A2.
XX
PD 15-NOV-2001.
XX
PF 10-MAY-2001; 2001WO-US015412.

XX
PR 10-MAY-2000; 2000US-0202529P.
XX
PA (PHYC-) PHYCOGEN INC.
XX
PI Alberte RS, Smith RD;
XX
DR WPI; 2002-121947/16.
XX
PT New transgenic plants comprising a zosteric acid biosynthetic gene, a
PT saline resistance gene or a hypoxia resistance gene derived from Zostera
PT marina, useful for producing plants with antifouling traits.
XX
PS Example; Fig 3; 117pp; English.
XX
CC The present invention relates to a new transgenic plant comprising a
CC heterologous gene derived from the marine eelgrass Zostera marina or at
CC least one heterologous nucleotide sequence encoding a zosteric acid
CC biosynthetic function, a saline-resistance function, or a anoxia-
CC resistance function. The invention describes the method of producing a
CC transgenic plant possessing an anti-fouling genetic trait by providing a
CC cDNA population derived from a marine vascular plant, isolating from the
CC cDNA population a nucleic acid species which hybridises to a nucleic acid
CC that encodes a sulfotransferase (ST), an alcohol dehydrogenase (ADH),
CC phenylalanine ammonia lyase (PAL) or a cinnamate-4-hydroxylase (CH), and
CC transforming a target host plant with the isolated nucleic acid. The
CC plant is useful in the genetic engineering of plant species having
CC desirable genetic traits such as antifouling traits, salt and anoxia
CC resistance, and pathogen defence strategy. The expression of such
CC biosynthetic enzymes are sufficient to support the production of zosteric
CC acid and other sulphated phenolic compounds in a target plant. AAS20856-
CC AAS20862 represent degenerate or gene-specific PCR primers for cloning Z.
CC marina sulfotransferase
XX
SQ Sequence 27 BP; 8 A; 6 C; 5 G; 8 T; 0 U; 0 Other;

Query Match 36.4%; Score 16; DB 6; Length 27;
Best Local Similarity 79.2%; Pred. No. 5.2e+03;
Matches 19; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 20 AATAACCGGTCGCGGTTATTAAGA 43
|||||| |||| ||||| | |||
Db 25 AATAACTTGTCGGGGTTAATCAGA 2

RESULT 4
AAS20860
ID AAS20860 standard; DNA; 27 BP.
XX
AC AAS20860;
XX
DT 09-APR-2002 (first entry)
XX
DE Gene-specific PCR primer Z-ST-P19 for cloning Z. marina sulfotransferase.
XX
KW Plant; transgenic; marine eelgrass; zosteric acid biosynthesis;
KW saline-resistance; anoxia-resistance; anti-fouling genetic trait;
KW marine vascular plant; sulphated phenolic compound; Zostera marina;
KW sulfotransferase; ST; enzyme; PCR; primer; ss.
XX
OS Zostera marina.
XX
PN WO200185971-A2.
XX
PD 15-NOV-2001.
XX
PF 10-MAY-2001; 2001WO-US015412.
XX
PR 10-MAY-2000; 2000US-0202529P.
XX
PA (PHYC-) PHYCOGEN INC.
XX
PI Alberte RS, Smith RD;
XX
DR WPI; 2002-121947/16.
XX
PT New transgenic plants comprising a zosteric acid biosynthetic gene, a
PT saline resistance gene or a hypoxia resistance gene derived from Zostera
PT marina, useful for producing plants with antifouling traits.
XX
PS Example; Fig 3; 117pp; English.
XX
CC The present invention relates to a new transgenic plant comprising a
CC heterologous gene derived from the marine eelgrass Zostera marina or at
CC least one heterologous nucleotide sequence encoding a zosteric acid
CC biosynthetic function, a saline-resistance function, or a anoxia-
CC resistance function. The invention describes the method of producing a
CC transgenic plant possessing an anti-fouling genetic trait by providing a
CC cDNA population derived from a marine vascular plant, isolating from the
CC cDNA population a nucleic acid species which hybridises to a nucleic acid
CC that encodes a sulfotransferase (ST), an alcohol dehydrogenase (ADH),
CC phenylalanine ammonia lyase (PAL) or a cinnamate-4-hydroxylase (CH), and
CC transforming a target host plant with the isolated nucleic acid. The
CC plant is useful in the genetic engineering of plant species having
CC desirable genetic traits such as antifouling traits, salt and anoxia
CC resistance, and pathogen defence strategy. The expression of such
CC biosynthetic enzymes are sufficient to support the production of zosteric
CC acid and other sulphated phenolic compounds in a target plant. AAS20856-
CC AAS20862 represent degenerate or gene-specific PCR primers for cloning Z.
CC marina sulfotransferase
XX
SQ Sequence 27 BP; 8 A; 5 C; 6 G; 8 T; 0 U; 0 Other;

Query Match 36.4%; Score 16; DB 6; Length 27;
Best Local Similarity 79.2%; Pred. No. 5.2e+03;
Matches 19; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 20 AATAACCGGTCGCGGTTATTAAGA 43
|||||| |||| ||||| | |||
Db 3 AATAACTTGTCGGGGTTAATCAGA 26

RESULT 5
ADE85746
ID ADE85746 standard; DNA; 30 BP.
XX
AC ADE85746;
XX
DT 12-FEB-2004 (first entry)
XX
DE EphA2 inverted antisense oligonucleotide control SEQ ID NO:50.
XX
KW cancer; hyperproliferative cell disease; EphA2 antibody;
KW EphA2 agonistic antibody; cytostatic; antiasthmatic; antipsoriatic;
KW antiinflammatory; vasotropic; respiratory; gene therapy;
KW metastatic cancer; asthma; psoriasis; inflammatory bowel disease;
KW smooth muscle restenosis; endothelial restenosis; Crohn's disease;
KW chronic obstructive pulmonary disease; human; control; ss.
XX
OS Synthetic.
OS Homo sapiens.
XX
PN WO2003094859-A2.
XX
PD 20-NOV-2003.
XX
PF 12-MAY-2003; 2003WO-US015044.
XX
PR 10-MAY-2002; 2002US-0379322P.
PR 14-OCT-2002; 2002US-0418213P.
PR 03-APR-2003; 2003US-0460507P.
XX
PA (MEDI-) MEDIMMUNE INC.
XX
PI Kinch MS, Carles-Kinch K, Kiener P, Langermann S;
XX

DR WPI; 2004-012002/01.
XX
PT Treating cancer or a non-cancer hyperproliferative cell disease (e.g.
PT asthma, psoriasis, inflammatory bowel disease or restenosis) in a patient
PT comprises administering to the patient a therapeutic amount of an EphA2
PT antibody.
XX
PS Example; SEQ ID NO 50; 173pp; English.
XX
CC The present invention describes a method for treating cancer or a non-
CC cancer hyperproliferative cell disease or disorder in a patient, which
CC comprises administering to the patient a therapeutic amount of an EphA2
CC antibody (I) that is an EphA2 agonistic antibody, an EphA2 cancer cell
CC phenotype inhibiting antibody, an exposed EphA2 epitope antibody, or an
CC antibody that binds EphA2 with a K-off of less than 3 x 10-3 s-1. Also
CC described: (1) a pharmaceutical composition comprising a therapeutic
CC amount of (I) and a pharmaceutical carrier; (2) a cell line that produces
CC (I); (3) a hybridoma deposited with the ATCC accession number PTA-4572,
CC PTA-4573 or PTA-4574; (4) an isolated nucleic acid comprising a
CC nucleotide sequence encoding a light chain variable domain or a heavy
CC chain variable domain of the EphA2 antibody; (5) a vector comprising the
CC nucleic acid described above; (6) a host cell comprising the vector; (7)
CC methods of identifying the EphA2 agonistic antibody or the EphA2 antibody
CC that inhibits a cancer cell phenotype, that kills cancer cells having a
CC cancer cell phenotype or that preferentially binds an EphA2 epitope
CC exposed on cancer cells; and (8) a method of diagnosing, prognosing or
CC monitoring the efficacy of therapy for cancer in a patient known to or
CC suspected to have cancer. (I) has cytostatic, antiasthmatic,
CC antipsoriatic, antiinflammatory, vasotropic and respiratory activities,
CC and can be used in gene therapy. The composition and methods are useful
CC in managing, diagnosing, preventing or treating hyperproliferative cell
CC diseases (i.e. metastatic cancer) or non-cancer hyperproliferative cell
CC diseases or disorders, such as asthma, psoriasis, inflammatory bowel
CC disease, smooth muscle restenosis, endothelial restenosis, Crohn's
CC disease or chronic obstructive pulmonary disease. They may also be used
CC for monitoring the efficacy of therapy for cancer in a patient known to
CC or suspected to have cancer, and in screening for anti-cancer drugs. The
CC present sequence is used in the exemplification of the present invention.
XX
SQ Sequence 30 BP; 4 A; 14 C; 6 G; 6 T; 0 U; 0 Other;

```
  Query Match             36.4%;  Score 16;  DB 12;  Length 30;
  Best Local Similarity   79.2%;  Pred. No. 5.3e+03;
  Matches   19;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          2 CGGGTCCCGTTCCTTCTTAATAAC 25
              || |||||||||||||   || ||
Db          3 CGCGTCCCGTTCCTTCACCATGAC 26
```

RESULT 6
ADL27336/c
ID ADL27336 standard; DNA; 33 BP.
XX
AC ADL27336;
XX
DT 03-JUN-2004 (first entry)
XX
DE Forward primer for amplifying prethrombin for cloning into pSecTag2A.
XX
KW adzyme; catalytic domain; allergic disease; inflammatory disease;
KW inflammatory disorder; allergic disorder; inflammatory cytokine; PCR;
KW primer; ss.
XX
OS Synthetic.
XX
PN WO2004019878-A2.
XX
PD 11-MAR-2004.
XX
PF 27-AUG-2003; 2003WO-US026937.
XX
PR 27-AUG-2002; 2002US-0406517P.
PR 05-NOV-2002; 2002US-0423754P.
PR 27-NOV-2002; 2002US-0430001P.
XX
PA (COMP-) COMPOUND THERAPEUTICS INC.
PA (AFEY/) AFEYAN N B.
XX
PI Afeyan NB, Baynes B, Dasgupta R, Lee FD, Wong GG;
XX
DR WPI; 2004-239110/22.
XX
PT New adzyme for enzymatically altering a substrate, useful for preparing a
PT composition for treating diseases associated with a soluble or membrane
PT bound molecule, e.g. allergic or inflammatory diseases.
XX
PS Example 2; Page 143; 202pp; English.
XX
CC The specification describes an adzyme for enzymatically altering a
CC substrate. The adzyme comprises a catalytic domain that catalyzes a
CC chemical reaction converting the substrate to one or more products, and a
CC targeting group that reversibly binds with an address site on the
CC substrate or with an address site on a second molecule that occurs in
CC functional proximity to the substrate (the targeting moiety and the
CC catalytic domain are heterologous with respect to each other). The adzyme
CC is useful for preparing a composition for treating diseases associated
CC with a soluble or membrane bound molecule, e.g., allergic or inflammatory
CC disease. Adzymes of the invention can be used to treat an inflammatory or
CC allergic disorder, where the substrate of the adzyme is an inflammatory
CC cytokine. PCR primers ADL27334-ADL27341 were used in overlap/recombinant
CC PCR to assemble prethrombin(G4S)3scFvalphaHA and scHA(G4S)3prethrombin,
CC which are adzymes of the invention.
XX
SQ Sequence 33 BP; 9 A; 11 C; 8 G; 5 T; 0 U; 0 Other;

```
  Query Match             36.4%;  Score 16;  DB 12;  Length 33;
  Best Local Similarity   79.2%;  Pred. No. 5.4e+03;
  Matches   19;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy         19 TAATAACCGGTCGCGGTTATTAAG 42
              || | || ||| ||||| ||||||
Db         32 TACTCACTGGTGGCGGTCATTAAG 9
```

RESULT 7
AAQ87806/c
ID AAQ87806 standard; DNA; 45 BP.
XX
AC AAQ87806;
XX
DT 25-MAR-2003 (revised)
DT 14-NOV-1995 (first entry)
XX
DE Primer used to amplify IgE receptor gene mRNA.
XX
KW Probe; immunoglobulin; IgE; receptor; beta; allergic disease; detection;
KW screening; diagnosis; ss.
XX
OS Synthetic.
XX
PN EP649910-A1.
XX
PD 26-APR-1995.
XX
PF 21-OCT-1994; 94EP-00307751.
XX
PR 22-OCT-1993; 93JP-00265144.
XX
PA (SUME) SUMITOMO ELECTRIC IND CO.
XX
PI Osoegawa M, Miyabe Y, Nakata M, Ra C, Suzuki K;
XX
DR WPI; 1995-156760/21.
XX
PT Probes for mutation(s) in beta chain gene of a high affinity IgE receptor

PT - for the diagnosis of allergic disease, esp. in neonate(s).
XX
PS Example 3; Page 9; 17pp; English.
XX
CC DNA probes (See AAQ87799-803) having sequences identical or complementary
CC to parts of the immunoglobulin E (IgE) receptor, are used to detect genes
CC associated with allergic disease (diseases involving mutations in the
CC beta chain gene). They may be used in neonatal screening, prenatal
CC diagnosis etc. Two primers (AAQ87805, AAQ87806) were used to amplify mRNA
CC in blood samples transcribed from the IgE receptor gene to produce cDNA.
CC The cDNA is then conjugated with a probe and an IgE ligand (See AAQ87804)
CC which is immobilised on a support is then used to bind to probe/IgE
CC receptor gene conjugates. The probe is then eluted by a gradual rise in
CC temperature. (Updated on 25-MAR-2003 to correct PN field.)
XX
SQ Sequence 45 BP; 13 A; 11 C; 11 G; 10 T; 0 U; 0 Other;

Query Match 35.9%; Score 15.8; DB 2; Length 45;
Best Local Similarity 74.1%; Pred. No. 6.9e+03;
Matches 20; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 6 TCCCGTTCCTTCTTAATAACCGGTCGC 32
|||| || || || |||| ||| |||
Db 31 TCCCATTGATTATTTATAAGCGGCCGC 5

RESULT 8
ABK89998/c
ID ABK89998 standard; DNA; 33 BP.
XX
AC ABK89998;
XX
DT 21-OCT-2002 (first entry)
XX
DE Human heavy chain CDR3 variable region, PCR primer eJH.
XX
KW Human; immune response; chronic B-lymphoproliferative disorder; CDR3;
KW complementarity determining region 3; hypervariable region; B-cell;
KW immunoglobulin heavy chain; VH-CDR3; idiotypic immunoglobulin;
KW cytostatic; PCR; primer; ss.
XX
OS Homo sapiens.
XX
PN WO200255559-A1.
XX
PD 18-JUL-2002.
XX
PF 15-JAN-2001; 2001WO-IT000014.
XX
PR 15-JAN-2001; 2001WO-IT000014.
XX
PA (FAZI/) FAZIO V M.
PA (SAGL/) SAGLIO G.
XX
PI Fazio VM, Saglio G;
XX
DR WPI; 2002-583654/62.
XX
PT Use of DNA sequences coding for hypervariable region (VH- complementarity
PT determining region 3 (CDR3)) of idiotypic immunoglobulin expressed on B-
PT cells of chronic B- lymphoproliferative disorders, as therapeutic
PT vaccine.
XX
PS Example 2; Fig 1B; 30pp; English.
XX
CC The present invention relates to a method for inducing an immune response
CC against B-lymphoproliferative disorders. The method comprises DNA
CC sequences encoding for the complementarity determining region 3 (CDR3)
CC hypervariable region of immunoglobulin heavy chain (VH-CDR3) alone or in
CC combination with at least another immunomodulating sequence. The DNA
CC sequences are useful as therapeutic vaccines for chronic B-
CC lymphoproliferative disorders in mammals, preferably humans. A
CC recombinant plasmid expression vector containing a DNA sequence of the
CC invention is useful as a therapeutic vaccine or for the manufacture of a
CC vaccine effective against chronic B-lymphoproliferative disorders
CC expressing the surface idiotypic immunoglobulin on B-cells in mammals,
CC preferably humans. An efficient, safe and easily reproducible DNA-based
CC immune response against B-lymphoproliferative pathologies can be
CC achieved. The present sequence represents a PCR primer used to amplify
CC human heavy chain CDR3 variable region in the examples of the present
CC invention
XX
SQ Sequence 33 BP; 7 A; 7 C; 11 G; 8 T; 0 U; 0 Other;

Query Match 35.5%; Score 15.6; DB 6; Length 33;
Best Local Similarity 70.0%; Pred. No. 8e+03;
Matches 21; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

Qy 3 GGGTCCCGTTCCTTCTTAATAACCGGTCGC 32
|| |||| | || ||||||| ||| |||
Db 33 GGTCACCGTCTCCTCATAATAAGCGGCCGC 4

RESULT 9
AAX27543
ID AAX27543 standard; DNA; 34 BP.
XX
AC AAX27543;
XX
DT 27-MAY-1999 (first entry)
XX
DE Staphylokinase (Sak) encoding DNA amplifying primer 2.
XX
KW Staphylokinase; Sak; recombinant; myocardial infarction; cerebral;
KW thrombembolia disease; arterial thrombosis; pulmonary thrombosis;
KW hydrolytic; fibrin; PCR primer; ss.
XX
OS Synthetic.
OS Staphylococcus aureus.
XX
PN WO9904017-A1.
XX
PD 28-JAN-1999.
XX
PF 17-JUL-1998; 98WO-CN000129.
XX
PR 19-JUL-1997; 97CN-00105988.
XX
PA (BAIH/) BAI H.
XX
PI Zhang Q, Zhang G, Xu G, Qu G, Bie L, Xu W, Wu Y;
XX
DR WPI; 1999-132261/11.
XX
PT Highly safe, novel recombinant staphylokinase (Sak) produced from high-
PT expression engineered strain - as plasminogen activator, with very high
PT hydrolytic activity to human fibrin, useful in treating thrombembolia
PT diseases e.g. myocardial infarction.
XX
PS Claim 12; Page 26; 51pp; Chinese.
XX
CC Sequences AAX27542-43 represent PCR primers for the amplification of the
CC DNA encoding a recombinant staphylokinase (Sak). The Sak-producing
CC Staphylococcus aureus SL1.063 is deposited as CGMCC No.0353. The
CC invention provides a method for constructing a Sak-producing engineered
CC strain which comprises (a) screening SaK-producing Staphylococcus aureus;
CC obtaining a required Sak gene by PCR (polymerase chain reaction)
CC amplification with chromosomal DNA of Sak-producing S. aureus as
CC template, with primers (AAX27542-43) (b) introduction of the obtained DNA
CC fragment into a plasmid selected from pUC19 and pBV220; and (c)
CC transferring the recombinant plasmid into a host cell such as E. coli of
CC DH5 alpha, TG1 or TG2 strains. The staphylokinase can be applied in
CC treatment of myocardial infarction, thrombembolia diseases and arterial
CC thrombosis including cerebral and pulmonary thrombi. Hydrolytic activity
CC of the staphylokinase to human fibrin is very high
XX

```
SQ   Sequence 34 BP; 8 A; 9 C; 4 G; 13 T; 0 U; 0 Other;

  Query Match             35.0%;  Score 15.4;  DB 2;  Length 34;
  Best Local Similarity   76.0%;  Pred. No. 9.8e+03;
  Matches   19;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          1 GCGGGTCCCGTTCCTTCTTAATAAC 25
              |||| |||  || ||| | ||||||
Db          2 GCGGATCCTATTTCTTTTCAATAAC 26


RESULT 10
AAF83844
ID   AAF83844 standard; DNA; 37 BP.
XX
AC   AAF83844;
XX
DT   11-SEP-2003  (revised)
DT   06-AUG-2001  (first entry)
XX
DE   C. pneumoniae amino acid transporter DNA amplifying 3' primer.
XX
KW   Amino acid transporter; Chlamydia; infection; chronic bronchitis;
KW   sinusitis; antibacterial; vaccine; gene therapy; PCR primer; ss.
XX
OS   Chlamydophila pneumoniae.
XX
PN   WO200136457-A2.
XX
PD   25-MAY-2001.
XX
PF   10-NOV-2000; 2000WO-CA001346.
XX
PR   15-NOV-1999;   99US-0165615P.
XX
PA   (AVET ) AVENTIS PASTEUR LTD.
XX
PI   Murdin AD,  Oomen RP,  Wang J,  Dunn P;
XX
DR   WPI; 2001-343797/36.
XX
PT   A Chlamydia polypeptide, an amino acid transporter gene, for the
PT   treatment and prevention of Chlamydia infection.
XX
PS   Claim 32; Page 52; 80pp; English.
XX
CC   Sequences AAF83843-844 represent PCR primers for amplifying the genomic
CC   DNA encoding a Chlamydia pneumoniae amino acid transporter protein. The
CC   amino acid transporter is useful for the treatment, prevention and
CC   diagnosis of Chlamydia infection, preferably Chlamydia pneumoniae
CC   infection, in human and veterinary applications. A protective vaccine
CC   against Chlamydia pneumoniae is useful to prevent infection which leads
CC   to chronic bronchitis and sinusitis. (Updated on 11-SEP-2003 to
CC   standardise OS field)
XX
SQ   Sequence 37 BP; 6 A; 10 C; 9 G; 12 T; 0 U; 0 Other;

  Query Match             34.5%;  Score 15.2;  DB 5;  Length 37;
  Best Local Similarity   71.4%;  Pred. No. 1.2e+04;
  Matches   20;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Qy          2 CGGGTCCCGTTCCTTCTTAATAACCGGT 29
              ||| ||||| || || |  |||| || |
Db          5 CGGATCCCGCTCTTTTTAGATAAGCGTT 32


RESULT 11
ADB81046/c
ID   ADB81046 standard; DNA; 38 BP.
XX
AC   ADB81046;
XX
DT   04-DEC-2003  (first entry)
XX
DE   LINE retro-position related primer, SEQ ID No 6.
XX
KW   RNA retro-position; 3' UTR; LINE; APE domain; retro-transposition;
KW   endonuclease domain; chromosome; gene therapy; gene transfer; PCR;
KW   primer; ss.
XX
OS   Unidentified.
XX
PN   WO2003064644-A1.
XX
PD   07-AUG-2003.
XX
PF   26-NOV-2002; 2002WO-JP012317.
XX
PR   31-JAN-2002; 2002JP-00024226.
XX
PA   (DNAV-) DNAVEC RES INC.
XX
PI   Fujiwara H,  Takahashi H,  Hasegawa M;
XX
DR   WPI; 2003-627609/59.
XX
PT   LINE retro-position by trans-complementation for transferring targeted,
PT   specific gene or nucleic acid of e.g. endonuclease domain via
PT   substitution to chromosome using virus vector, applicable in gene
PT   therapy.
XX
PS   Example 1; Page 35; 96pp; Japanese.
XX
CC   The invention relates to a novel RNA retro-position comprising the
CC   transcription of an RNA containing a 3' UTR fragment of a LINE in cells;
CC   and trans-positioning the ORF protein of such LINE after expressing from
CC   other than the RNA. The invention further comprises a similar method in
CC   which the transcription of an RNA containing a 3' UTR fragment of an APE
CC   domain-carrying type site-specific LINE in cells, and expressing the ORF
CC   protein of the LINE in such cells; or transcription of an RNA containing
CC   3' UTR fragment of a LINE in cells, and expressing ORF protein in such
CC   cells thereby modifying a retro-transposition target site of a LINE by
CC   substituting the endonuclease domain of the LINE by that of another LINE
CC   via ORF protein of such LINE. The invention also includes a retro-
CC   transposition vector with RNA encoding the 3' UTR fragment of a LINE but
CC   not expressing the encoded ORF of the LINE; a vector encoding a protein
CC   for substitution of the endonuclease domain of an encoded ORF protein in
CC   the site-specific LINE by the endonuclease domain of the encoded ORF
CC   protein in another LINE; and a kit for gene transfer through retro-
CC   transposition of an RNA. The method is useful for transferring targeted,
CC   specific genes or nucleic acids of an endonuclease domain via
CC   substitution to a chromosome using a virus vector, which is applicable in
CC   gene therapy. The retro-transposition in the host is highly efficient by
CC   targeting specifically at LINE, and with little damage to the host due to
CC   the gene transfer. This polynucleotide sequence represents a PCR primer
CC   used in the exemplification of the invention.
XX
SQ   Sequence 38 BP; 18 A; 7 C; 7 G; 6 T; 0 U; 0 Other;

  Query Match             34.5%;  Score 15.2;  DB 9;  Length 38;
  Best Local Similarity   71.4%;  Pred. No. 1.2e+04;
  Matches   20;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Qy         12 TCCTTCTTAATAACCGGTCGCGGTTATT 39
              | |||||| ||||| |  |  |||| ||
Db         29 TTCTTCTTTATAACTGCCCATGGTTTTT 2


RESULT 12
AAZ65893/c
ID   AAZ65893 standard; DNA; 47 BP.
XX
AC   AAZ65893;
XX
DT   11-SEP-2001  (first entry)
XX
```

DE Human map-related biallelic marker SEQ ID NO:240.
XX
KW Human genome; biallelic marker; high density disequilibrium map;
KW genomic map; haplotype; phenotype; polymorphic base; genotyping;
KW haplotyping; hybridisation; identification; characterisation; diagnosis;
KW single nucleotide polymorphism; SNP; ds.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT variation replace(24,T)
FT /*tag= a
FT /standard_name= "single nucleotide polymorphism"
XX
PN WO9954500-A2.
XX
PD 28-OCT-1999.
XX
PF 21-APR-1999; 99WO-IB000822.
XX
PR 21-APR-1998; 98US-0082614P.
PR 23-NOV-1998; 98US-0109732P.
XX
PA (GEST) GENSET.
XX
PI Cohen D, Blumenfeld M, Chumakov I;
XX
DR WPI; 2000-013267/01.
XX
PT Novel biallelic markers used to construct a high density disequilibrium
PT map of the human genome.
XX
PS Claim 1; Page 281; 2745pp; English.
XX
CC AAZ65654 to AAZ69578 represent human biallelic markers from the present
CC invention, which contain a polymorphic base at position 24 of their
CC nucleotide sequences. AAZ69579 to AAZ77440 represent amplification
CC primers for the biallelic markers. The biallelic markers of the invention
CC have a variety of uses: they can be used for high density mapping of the
CC human genome, and in complex association studies and haplotyping studies
CC which are useful in determining the genetic basis for disease states.
CC Compositions and methods of the invention can also be useful for the
CC identification of the targets for the development of pharmaceutical
CC agents and diagnostic methods, as well as the characterisation of the
CC differential efficacious responses to and side effects from
CC pharmaceutical agents acting on a disease as well as other treatment.
CC N.B. The SEQ ID NOs 2852, 2913, 2974, 3035, 3096, 3157, 3227, 3297 and
CC 3367, are not actually given a sequence in the Sequence Listing from the
CC present invention
XX
SQ Sequence 47 BP; 18 A; 6 C; 7 G; 16 T; 0 U; 0 Other;

Query Match 34.5%; Score 15.2; DB 3; Length 47;
Best Local Similarity 63.9%; Pred. No. 1.3e+04;
Matches 23; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

Qy 9 CGTTCCTTCTTAATAACCGGTCGCGGTTATTAAGAA 44
| || | |||||| | |||| ||| | || ||
Db 46 CATTTAATTTTAATACATGCTCGCTGTTTTGAAAAA 11

RESULT 13
AAF91472/c
ID AAF91472 standard; DNA; 47 BP.
XX
AC AAF91472;
XX
DT 04-MAY-2001 (first entry)
XX
DE N. meningitidis NspA upstream sequence inverse PCR primer PNS4.
XX
KW Modified Gram-negative bacterium; outer membrane vesicle; bleb; vaccine;
KW genetically modified; protective antigen expression; LPS detoxification;
KW LPS; lipid A; homologous recombination vector; immunisation;
KW immunoprotective; non-toxic; paediatric; plasmid construction;
KW modified Neisseria meningitidis; PCR primer; ss.
XX
OS Neisseria meningitidis.
XX
PN WO200109350-A2.
XX
PD 08-FEB-2001.
XX
PF 31-JUL-2000; 2000WO-EP007424.
XX
PR 03-AUG-1999; 99GB-00018319.
XX
PA (SMIK) SMITHKLINE BEECHAM BIOLOGICALS.
XX
PI Berthet FJ, Dalemans WLJ, Denoel P, Dequesne G, Feron C, Lobet Y;
PI Poolman J, Thiry G, Thonnard J, Voet P;
XX
DR WPI; 2001-138654/14.
XX
PT New isolated polynucleotide useful for outer membrane vesicle preparation
PT from Gram-negative bacterial strain for vaccination of microbial
PT infections.
XX
PS Example 6; Page 47; 128pp; English.
XX
CC The invention relates to a genetically-engineered outer membrane vesicle
CC (bleb) preparation from a Gram-negative bacterium for use as a vaccine.
CC The blebs of the invention are improved with respect to their
CC immunogenicity and toxicity by the introduction of one or more genetic
CC changes to the chromosome of the bacterium from which the blebs are
CC derived. The changes made include the upregulation of protective antigen
CC expression, the downregulation of immunodominant non-protective antigen
CC expression, and genetic changes which result in detoxification of the
CC Lipid A moiety of lipopolysaccharide (LPS). The invention also
CC encompasses modified Gram-negative bacterial strains from which the bleb
CC preparations are made, a vector suitable for performing recombination
CC events (for the generation of the modified bacterial strains),
CC bacterially-derived nucleic acid sequences used in such a vector, and an
CC immunoprotective and non-toxic Gram-negative bleb, ghost, or killed whole
CC cell vaccine suitable for paediatric use. The bleb preparation is useful
CC in the manufacture of a medicament for immunising a human host against a
CC disease caused by infection of one or more of the following: Neisseria
CC meningitidis, Neisseria gonorrhoeae, Haemophilus influenza, Moraxella
CC catarrhalis, Pseudomonas aeruginosa, Chlamydia trachomatis, and Chlamydia
CC pneumonia. The invention may also be used to provide immunisation against
CC the influenza virus. Bacterially derived nucleotide sequences of the
CC invention are used in the performance of homologous recombination events
CC up to 1000 bp upstream of a bacterial chromosomal gene in order to either
CC increase or decrease expression of that gene. Immunoprotective and non-
CC toxic Gram-negative bleb, ghost, or killed whole cell vaccines are more
CC immunogenic, less toxic and safer, and are particularly useful for
CC paediatric use. The present sequence represents a PCR primer used in the
CC production of a plasmid used in the generation of a modified Neisseria
CC meningitidis strain
XX
SQ Sequence 47 BP; 15 A; 9 C; 6 G; 17 T; 0 U; 0 Other;

Query Match 34.5%; Score 15.2; DB 4; Length 47;
Best Local Similarity 71.4%; Pred. No. 1.3e+04;
Matches 20; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 17 CTTAATAACCGGTCGCGGTTATTAAGAA 44
| || | |||| |||| | |||||| |
Db 41 CATATTTTCCGGACGCGTTAATTAAGGA 14

RESULT 14
ABK37852/c
ID ABK37852 standard; DNA; 47 BP.
XX
AC ABK37852;

```
XX
DT   08-MAY-2002  (first entry)
XX
DE   pCMK(+) construction PCR primer #17.
XX
KW   pCMK(+); ss; PCR; primer; antibacterial; vaccine; bleb;
KW   Gram-negative bacteria; outer membrane; LPS; lipopolysaccharide;
KW   meningitis; bacteraemia; otitis media; pneumonia; chronic bronchitis;
KW   sinusitis.
XX
OS   Neisseria meningitidis.
XX
PN   WO200209746-A2.
XX
PD   07-FEB-2002.
XX
PF   31-JUL-2001; 2001WO-EP008857.
XX
PR   31-JUL-2000; 2000WO-EP007424.
PR   08-FEB-2001; 2001GB-00003170.
XX
PA   (SMIK ) SMITHKLINE BEECHAM BIOLOGICALS.
XX
PI   Berthet FJ,  Dalemans W,  Denoel P,  Dequesne G,  Feron C,  Garcon N;
PI   Lobet Y,  Poolman J,  Thiry G,  Thonnard J,  Voet P;
XX
DR   WPI; 2002-188688/24.
XX
PT   New immunogenic composition comprising an antigen derived from a pathogen
PT   and a blep preparation from Neisserria meningitidis, useful as a vaccine
PT   for treating or preventing disease caused by the pathogen.
XX
PS   Example 2; Page 49; 125pp; English.
XX
CC   The invention relates to an immunogenic composition comprising an antigen
CC   derived from a pathogen capable of protecting a host against the
CC   pathogen, mixed with an adjuvant comprising a bleb preparation derived
CC   from a Gram-negative bacterial strain. The immunogenic composition
CC   consists of N. meningitidis B blebs or N. meningitidis C polysaccharide
CC   antigen. The blebs (derived from the outer membrane) may also have their
CC   toxic lipopolysaccharide (LPS) content reduced using heterologous down
CC   regulating sequences for LPS pathway genes or by up regulating genes
CC   involved in LPS synthesis suppression, by a promoter replacement
CC   technique. The immunogenic preparation is useful in the manufacture of a
CC   medicament for the treatment of a disease caused by the pathogen from
CC   which the antigen is derived (e.g. from Neisseria, meningitis and
CC   bacteraemia, from Moraxella, otitis media and pneumonia, and from H.
CC   influenzae chronic bronchitis, sinusitis, pneumonia and otitis media).
CC   The bleb derived from M. catarrhalis or from a non-typeable H. influenzae
CC   is useful as an adjuvant in an immunogenic composition comprising one or
CC   more pneumococcal capsular polysaccharides or protein antigens. The
CC   present sequence is a PCR primer used to construct the gene delivery
CC   vector pCMK(+) for producing N. meningitidis strains producing
CC   recombinant blebs
XX
SQ   Sequence 47 BP; 15 A; 9 C; 6 G; 17 T; 0 U; 0 Other;

  Query Match             34.5%;  Score 15.2;  DB 6;  Length 47;
  Best Local Similarity   71.4%;  Pred. No. 1.3e+04;
  Matches   20;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Qy         17 CTTAATAACCGGTCGCGGTTATTAAGAA 44
              | ||  |  |||| |||| | |||||| |
Db         41 CATATTTTCCGGACGCGTTAATTAAGGA 14


RESULT 15
AAI30035/c
ID   AAI30035 standard; DNA; 31 BP.
XX
AC   AAI30035;
XX
DT   04-NOV-2004  (revised)
DT   18-OCT-2001  (first entry)
XX
DE   Human single nucleotide polymorphism (SNP) 72.
XX
KW   Human; resequence; genotype; disease; forensic; paternity testing;
KW   single nucleotide polymorphism; SNP; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   variation       16
FT                   /*tag=  a
FT                   /standard_name= "single nucleotide polymorphism"
XX
PN   WO200166800-A2.
XX
PD   13-SEP-2001.
XX
PF   07-MAR-2001; 2001WO-US007268.
XX
PR   07-MAR-2000; 2000US-0187510P.
PR   22-MAY-2000; 2000US-0206129P.
XX
PA   (WHED ) WHITEHEAD INST BIOMEDICAL RES.
XX
PI   Cargill M,  Ireland JS,  Lander ES;
XX
DR   WPI; 2001-522952/57.
XX
PT   Nucleic acid molecules from the human genome which include polymorphic
PT   sites, useful in methods for predicting the presence, absence or severity
PT   of a particular phenotype or disorder (e.g. diabetes) associated with a
PT   particular genotype.
XX
PS   Claim 1; Page 64; 145pp; English.
XX
CC   The invention relates to the identification of nucleic acid molecules
CC   (AAI29513-AAI31314) from the human genome which include polymorphic sites
CC   which can predispose individuals to disease. Various genes from a number
CC   of individuals were resequenced and single nucleotide polymorphisms
CC   (SNPs) in these genes discovered. The method is useful for predicting the
CC   presence, absence or severity of a particular phenotype or disorder (e.g.
CC   diabetes) associated with a particular genotype. The nucleic acids
CC   containing the polymorphic sites may be useful in forensics and paternity
CC   testing
CC
CC   Revised record issued on 04-NOV-2004 : Correction to Feature Table Key
XX
SQ   Sequence 31 BP; 12 A; 7 C; 9 G; 3 T; 0 U; 0 Other;

  Query Match             34.1%;  Score 15;  DB 4;  Length 31;
  Best Local Similarity   78.3%;  Pred. No. 1.4e+04;
  Matches   18;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy         12 TCCTTCTTAATAACCGGTCGCGG 34
              || |||||||| | | |||| ||
Db         29 TCTTTCTTAATGAGCTGTCGGGG 7


RESULT 16
AAZ69059
ID   AAZ69059 standard; DNA; 47 BP.
XX
AC   AAZ69059;
XX
DT   10-SEP-2001  (first entry)
XX
DE   Human map-related biallelic marker SEQ ID NO:3415.
XX
KW   Human genome; biallelic marker; high density disequilibrium map;
KW   genomic map; haplotype; phenotype; polymorphic base; genotyping;
KW   haplotyping; hybridisation; identification; characterisation; diagnosis;
KW   single nucleotide polymorphism; SNP; ds.
```

```
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   variation       replace(24,G)
FT                   /*tag=  a
FT                   /standard_name= "single nucleotide polymorphism"
XX
PN   WO9954500-A2.
XX
PD   28-OCT-1999.
XX
PF   21-APR-1999;   99WO-IB000822.
XX
PR   21-APR-1998;   98US-0082614P.
PR   23-NOV-1998;   98US-0109732P.
XX
PA   (GEST ) GENSET.
XX
PI   Cohen D,  Blumenfeld M,  Chumakov I;
XX
DR   WPI; 2000-013267/01.
XX
PT   Novel biallelic markers used to construct a high density disequilibrium
PT   map of the human genome.
XX
PS   Claim 3; Page 959; 2745pp; English.
XX
CC   AAZ65654 to AAZ69578 represent human biallelic markers from the present
CC   invention, which contain a polymorphic base at position 24 of their
CC   nucleotide sequences. AAZ69579 to AAZ77440 represent amplification
CC   primers for the biallelic markers. The biallelic markers of the invention
CC   have a variety of uses: they can be used for high density mapping of the
CC   human genome, and in complex association studies and haplotyping studies
CC   which are useful in determining the genetic basis for disease states.
CC   Compositions and methods of the invention can also be useful for the
CC   identification of the targets for the development of pharmaceutical
CC   agents and diagnostic methods, as well as the characterisation of the
CC   differential efficacious responses to and side effects from
CC   pharmaceutical agents acting on a disease as well as other treatment.
CC   N.B. The SEQ ID NOs 2852, 2913, 2974, 3035, 3096, 3157, 3227, 3297 and
CC   3367, are not actually given a sequence in the Sequence Listing from the
CC   present invention
XX
SQ   Sequence 47 BP; 16 A; 13 C; 5 G; 13 T; 0 U; 0 Other;

  Query Match             34.1%;  Score 15;  DB 3;  Length 47;
  Best Local Similarity   67.7%;  Pred. No. 1.6e+04;
  Matches   21;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy         10 GTTCCTTCTTAATAACCGGTCGCGGTTATTA 40
              |||| | ||||| ||||  || | | |  ||
Db         16 GTTCATCCTTAAAAACCATTCTCAGCTCCTA 46


RESULT 17
ABZ02614
ID   ABZ02614 standard; DNA; 50 BP.
XX
AC   ABZ02614;
XX
DT   09-JAN-2003  (first entry)
XX
DE   Human leukocyte gene expression profiling probe SEQ ID NO 2605.
XX
KW   T7; leukocyte; gene expression profiling; allograft rejection;
KW   atherosclerosis; congestive heart failure; systemic lupus erythematosus;
KW   rheumatoid arthritis; osteoarthritis; cytomegalovirus; infection; probe;
KW   ss.
XX
OS   Homo sapiens.
XX
PN   WO200257414-A2.
XX
PD   25-JUL-2002.
XX
PF   22-OCT-2001; 2001WO-US047856.
XX
PR   20-OCT-2000; 2000US-0241994P.
PR   08-JUN-2001; 2001US-0296764P.
XX
PA   (BIOC-) BIOCARDIA INC.
XX
PI   Wohlgemuth J,  Fry K,  Matcuk G,  Altman P,  Prentice J,  Phillips J;
PI   Ly N,  Woodward R,  Quertermous T,  Johnson F;
XX
DR   WPI; 2002-636525/68.
XX
PT   New system for leukocyte expression profiling, diagnosing a disease, or
PT   monitoring (the rate of) progression of a disease, e.g. atherosclerosis
PT   or congestive heart failure, comprises diagnostic oligonucleotides.
XX
PS   Claim 1; Page 410; 0pp; English.
XX
CC   The invention relates to a system for detecting gene expression, which
CC   comprises one or two isolated DNA molecules that detect expression of a
CC   gene, where the gene corresponds to any of 8143 oligonucleotides
CC   (ABZ00010-ABZ08152) each having 50 base pairs (bp). The system is useful
CC   for leukocyte expression profiling. It is particularly useful for
CC   diagnosing a disease, monitoring (rate of) progression of a disease,
CC   predicting therapeutic outcome, determining prognosis for a patient,
CC   predicting disease complications in an individual or monitoring response
CC   to treatment in an individual. The diseases include cardiac allograft
CC   rejection, kidney allograft rejection, liver allograft rejection,
CC   atherosclerosis, congestive heart failure, systemic lupus erythematosus,
CC   rheumatoid arthritis, osteoarthritis or cytomegalovirus infection
XX
SQ   Sequence 50 BP; 9 A; 8 C; 8 G; 25 T; 0 U; 0 Other;

  Query Match             34.1%;  Score 15;  DB 6;  Length 50;
  Best Local Similarity   67.7%;  Pred. No. 1.6e+04;
  Matches   21;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy         14 CTTCTTAATAACCGGTCGCGGTTATTAAGAA 44
              || ||| ||  ||  | ||| |||||  |||
Db          2 CTGCTTCATCTCCTTTTGCGCTTATTTGGAA 32


RESULT 18
ABZ07725/c
ID   ABZ07725 standard; DNA; 50 BP.
XX
AC   ABZ07725;
XX
DT   09-JAN-2003  (first entry)
XX
DE   Human leukocyte gene expression profiling probe SEQ ID NO 7716.
XX
KW   T7; leukocyte; gene expression profiling; allograft rejection;
KW   atherosclerosis; congestive heart failure; systemic lupus erythematosus;
KW   rheumatoid arthritis; osteoarthritis; cytomegalovirus; infection; probe;
KW   ss.
XX
OS   Homo sapiens.
XX
PN   WO200257414-A2.
XX
PD   25-JUL-2002.
XX
PF   22-OCT-2001; 2001WO-US047856.
XX
PR   20-OCT-2000; 2000US-0241994P.
PR   08-JUN-2001; 2001US-0296764P.
XX
PA   (BIOC-) BIOCARDIA INC.
XX
```

PI Wohlgemuth J, Fry K, Matcuk G, Altman P, Prentice J, Phillips J;
PI Ly N, Woodward R, Quertermous T, Johnson F;
XX
DR WPI; 2002-636525/68.
XX
PT New system for leukocyte expression profiling, diagnosing a disease, or
PT monitoring (the rate of) progression of a disease, e.g. atherosclerosis
PT or congestive heart failure, comprises diagnostic oligonucleotides.
XX
PS Claim 1; Page 576; 0pp; English.
XX
CC The invention relates to a system for detecting gene expression, which
CC comprises one or two isolated DNA molecules that detect expression of a
CC gene, where the gene corresponds to any of 8143 oligonucleotides
CC (ABZ00010-ABZ08152) each having 50 base pairs (bp). The system is useful
CC for leukocyte expression profiling. It is particularly useful for
CC diagnosing a disease, monitoring (rate of) progression of a disease,
CC predicting therapeutic outcome, determining prognosis for a patient,
CC predicting disease complications in an individual or monitoring response
CC to treatment in an individual. The diseases include cardiac allograft
CC rejection, kidney allograft rejection, liver allograft rejection,
CC atherosclerosis, congestive heart failure, systemic lupus erythematosus,
CC rheumatoid arthritis, osteoarthritis or cytomegalovirus infection
XX
SQ Sequence 50 BP; 15 A; 9 C; 13 G; 13 T; 0 U; 0 Other;

Query Match 34.1%; Score 15; DB 6; Length 50;
Best Local Similarity 67.7%; Pred. No. 1.6e+04;
Matches 21; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Qy 12 TCCTTCTTAATAACCGGTCGCGGTTATTAAG 42
| |||||| |||| ||||| || |||
Db 49 TTCTTCTTCATAATGAGTCGCTTTTGAAAAG 19

RESULT 19
ABK59783/c
ID ABK59783 standard; DNA; 31 BP.
XX
AC ABK59783;
XX
DT 02-JUL-2002 (first entry)
XX
DE Human CLCA1 gene enzymatic nucleic acid #4154.
XX
KW Human; chloride channel calcium activated 1; CLCA1; ss; antiasthmatic;
KW antiinflammatory; chronic obstructive pulmonary disease; COPD; asthma;
KW chronic bronchitis; cystic fibrosis; obstructive bowel syndrome;
KW oxygen therapy; bronchodilator; corticosteroid; vaccination; mucokinetic;
KW acetylcysteine.
XX
OS Homo sapiens.
XX
PN WO200211674-A2.
XX
PD 14-FEB-2002.
XX
PF 09-AUG-2001; 2001WO-US024970.
XX
PR 09-AUG-2000; 2000US-0224383P.
XX
PA (RIBO-) RIBOZYME PHARM INC.
PA (SYNT) SYNTEX USA LLC.
PA (THOM/) THOMPSON J.
XX
PI Thompson J, Mcswiggen J, Mckenzie T, Ayers D, Szymkowski DE;
PI Grupe A;
XX
DR WPI; 2002-217145/27.
XX
PT Enzymatic polynucleotide that down regulates expression of chloride
PT channel calcium activated gene, useful for treating Chronic obstructive
PT pulmonary disease (COPD), chronic bronchitis and asthma.
XX
PS Claim 5; Page 105; 152pp; English.
XX
CC The invention relates to enzymatic nucleic acid molecules that down
CC regulate expression of chloride channel calcium activated 1 (CLCA1) genes
CC by cleaving RNA derived from the genes. The nucleic acid sequences are
CC useful as pharmaceutical agents for treating conditions such as chronic
CC obstructive pulmonary disease (COPD), chronic bronchitis, asthma, cystic
CC fibrosis, obstructive bowel syndrome and any other diseases or conditions
CC that are related to or will respond to the levels of CLCA1 in a cell or
CC tissue. The sequences are useful for reducing CLCA1 activity in a cell,
CC hence, are useful for treatment of a patient having a condition
CC associated with the level of CLCA1, where the invention further comprises
CC the use of one or more therapies under conditions suitable for the
CC treatment, for example, oxygen therapy, bronchodilators, corticosteroids,
CC antibacterials, vaccinations, acetylcysteine and mucokinetic agents. The
CC nucleic acids of the invention are also used as diagnostic tools to
CC examine genetic drift and mutations within diseased cells or to detect
CC the presence of CLCA1 RNA in a cell. This sequence represents an
CC enzymatic nucleic acid molecule of the invention
XX
SQ Sequence 31 BP; 9 A; 7 C; 11 G; 4 T; 0 U; 0 Other;

Query Match 33.6%; Score 14.8; DB 6; Length 31;
Best Local Similarity 73.1%; Pred. No. 1.8e+04;
Matches 19; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 5 GTCCCGTTCCTTCTTAATAACCGGTC 30
||||||||| || | || || |||
Db 30 GTCCCGTTCGTTGTAGCTAGCCCGTC 5

RESULT 20
AAF81443
ID AAF81443 standard; DNA; 36 BP.
XX
AC AAF81443;
XX
DT 08-JUN-2001 (first entry)
XX
DE PCR primer GSP2 for corn promoter clone #700343485.
XX
KW Corn; promoter; transgenic plant; herbicide resistance; PCR primer; ss.
XX
OS Zea mays.
XX
PN WO200119976-A2.
XX
PD 22-MAR-2001.
XX
PF 13-SEP-2000; 2000WO-US025078.
XX
PR 16-SEP-1999; 99US-0154182P.
XX
PA (MONS) MONSANTO CO.
XX
PI Anderson HM, Chay CA, Chen G, Conner TW;
XX
DR WPI; 2001-244796/25.
XX
PT Novel promoter nucleic acid sequences useful for regulating heterologous
PT gene expression in plants, comprising regulatory sequences located
PT upstream to plant DNA structural coding sequences.
XX
PS Example 3; Page 87; 101pp; English.
XX
CC The present invention relates to novel corn promoter sequences (see
CC AAF81456-AAF81478). The promoter sequences are useful for conferring
CC expression of a second polynucleotide molecule in a transgenic plant
CC tissue. In addition, the promoter sequences are useful for providing
CC plants with herbicide resistance. The promoter sequences are suitable for
CC selectively modulating expression of any operatively linked gene and
CC provide additional regulatory element diversity in a plant expression

```
CC   vector in gene stacking approaches. The present sequence is a PCR primer
CC   used in the present invention
XX
SQ   Sequence 36 BP; 6 A; 11 C; 9 G; 10 T; 0 U; 0 Other;

  Query Match             33.6%;  Score 14.8;  DB 4;  Length 36;
  Best Local Similarity   73.1%;  Pred. No. 1.8e+04;
  Matches   19;  Conservative    0;  Mismatches    7;  Indels    0;  Gaps    0;

Qy          9 CGTTCCTTCTTAATAACCGGTCGCGG 34
              | || ||||| | | | |||| ||||
Db          6 CCTTTCTTCTCACTCAGCGGTTGCGG 31


RESULT 21
AAD43960/c
ID   AAD43960 standard; DNA; 39 BP.
XX
AC   AAD43960;
XX
DT   14-NOV-2002  (first entry)
XX
DE   DI228 oligo used to assemble beta-amyloid gene subfragment.
XX
KW   Amyloidogenic protein; Alzheimer's disease; Huntington's disease;
KW   spongiform encephalopathy; familial amyloid cardiomyopathy; amyloidosis;
KW   Gerstmann-Straussler-Scheinker syndrome; spongiform encephalopathy; GSS;
KW   Creutzfeldt-Jacob disease; insulinoma; diabetes; body myocytis; myeloma;
KW   CJ; beta-amyloid; ss.
XX
OS   Unidentified.
XX
PN   WO200242462-A2.
XX
PD   30-MAY-2002.
XX
PF   27-NOV-2001; 2001WO-US044581.
XX
PR   27-NOV-2000; 2000US-0253302P.
PR   29-NOV-2000; 2000US-0250198P.
PR   20-DEC-2000; 2000US-0257186P.
XX
PA   (PRAE-) PRAECIS PHARM INC.
XX
PI   Gefter ML,  Israel DI,  Joyal JL,  Gosselin M;
XX
DR   WPI; 2002-636427/68.
XX
PT   Novel therapeutic agent useful for treating an amyloidogenic disorder,
PT   e.g. Alzheimer's disease, comprises an immunoglobulin heavy chain
PT   constant region linked to a peptide capable of binding amyloidogenic
PT   protein.
XX
PS   Example 4; Fig 7A; 79pp; English.
XX
CC   The invention relates to a compound comprising an immunoglobulin (Ig)
CC   heavy chain constant region or its fragment that retains the ability to
CC   bind an Fc receptor linked by a linker group or a direct bond to a
CC   peptide capable of binding an amyloidogenic protein. The invention is
CC   useful for clearing an amyloidogenic protein such as beta-amyloid,
CC   transthyretin (TTR), prion protein (PrP), islet amyloid polypeptide
CC   (IAPP), atrial natriuretic factor (ANF), kappa light chain, lambda light
CC   chain, amyloid A, procalcitonin, cystatin C, beta2-microglobulin, ApoA-I,
CC   gelsolin, calcitonin, fibrinogen, Huntington, alpha-synuclein and
CC   lysozyme from a subject and for treating an amyloidogenic disorder such
CC   as Alzheimer's disease and spongiform encephalopathy. Disorders treatable
CC   include those caused or characterised by deposits of TTR (eg. familial
CC   amyloid cardiomyopathy), PrP (eg. spongiform encephalopathies, including
CC   scrapie in sheep, bovine spongiform encephalopathy in cows and
CC   Creutzfeldt-Jacob disease (CJ) and Gerstmann-Straussler-Scheinker
CC   syndrome (GSS) in humans), IAPP (eg. insulinoma, adult onset diabetes),
CC   ANF (eg. isolated atrial amyloid), kappa or lambda light chain (eg.
CC   idiopathic amyloidosis, myeloma), amyloid A (eg. amyloidosis), Apo A-I
CC   (eg. hereditary non-neuropathic systemic amyloidosis), Gelsolin (eg.
CC   familial amyloidosis of Finnish type), Fibrinogen (eg. hereditary renal
CC   amyloidosis), Lysozyme (eg. hereditary systemic amyloidosis). Other
CC   examples of amyloidogenic disorders include Huntington's disease and
CC   inclusion body myocytis. The present sequence is an oligonucleotide used
CC   to assemble beta amyloid gene subfragment
XX
SQ   Sequence 39 BP; 14 A; 10 C; 8 G; 7 T; 0 U; 0 Other;

  Query Match             33.6%;  Score 14.8;  DB 6;  Length 39;
  Best Local Similarity   73.1%;  Pred. No. 1.8e+04;
  Matches   19;  Conservative    0;  Mismatches    7;  Indels    0;  Gaps    0;

Qy          9 CGTTCCTTCTTAATAACCGGTCGCGG 34
              | |||||||||||    | ||| ||||
Db         36 CATTCCTTCTTATGTTCTGGTGGCGG 11


RESULT 22
AAZ67378/c
ID   AAZ67378 standard; DNA; 47 BP.
XX
AC   AAZ67378;
XX
DT   10-SEP-2001  (first entry)
XX
DE   Human map-related biallelic marker SEQ ID NO:1725.
XX
KW   Human genome; biallelic marker; high density disequilibrium map;
KW   genomic map; haplotype; phenotype; polymorphic base; genotyping;
KW   haplotyping; hybridisation; identification; characterisation; diagnosis;
KW   single nucleotide polymorphism; SNP; ds.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   variation       replace(24,T)
FT                   /*tag=  a
FT                   /standard_name= "single nucleotide polymorphism"
XX
PN   WO9954500-A2.
XX
PD   28-OCT-1999.
XX
PF   21-APR-1999;   99WO-IB000822.
XX
PR   21-APR-1998;   98US-0082614P.
PR   23-NOV-1998;   98US-0109732P.
XX
PA   (GEST ) GENSET.
XX
PI   Cohen D,  Blumenfeld M,  Chumakov I;
XX
DR   WPI; 2000-013267/01.
XX
PT   Novel biallelic markers used to construct a high density disequilibrium
PT   map of the human genome.
XX
PS   Claim 1; Page 598; 2745pp; English.
XX
CC   AAZ65654 to AAZ69578 represent human biallelic markers from the present
CC   invention, which contain a polymorphic base at position 24 of their
CC   nucleotide sequences. AAZ69579 to AAZ77440 represent amplification
CC   primers for the biallelic markers. The biallelic markers of the invention
CC   have a variety of uses: they can be used for high density mapping of the
CC   human genome, and in complex association studies and haplotyping studies
CC   which are useful in determining the genetic basis for disease states.
CC   Compositions and methods of the invention can also be useful for the
CC   identification of the targets for the development of pharmaceutical
CC   agents and diagnostic methods, as well as the characterisation of the
CC   differential efficacious responses to and side effects from
CC   pharmaceutical agents acting on a disease as well as other treatment.
CC   N.B. The SEQ ID NOs 2852, 2913, 2974, 3035, 3096, 3157, 3227, 3297 and
```

```
CC   3367, are not actually given a sequence in the Sequence Listing from the
CC   present invention
XX
SQ   Sequence 47 BP; 18 A; 4 C; 13 G; 12 T; 0 U; 0 Other;

  Query Match             33.6%;  Score 14.8;  DB 3;  Length 47;
  Best Local Similarity   59.5%;  Pred. No. 1.9e+04;
  Matches   25;  Conservative    0;  Mismatches   17;  Indels    0;  Gaps    0;

Qy          3 GGGTCCCGTTCCTTCTTAATAACCGGTCGCGGTTATTAAGAA 44
              || ||||  ||  |||| ||||     | |  ||| | | ||
Db         46 GGCTCCCCATCTCTCTTTATAAGTTAGCACCATTAATTATAA 5


RESULT 23
AAQ69835/c
ID   AAQ69835 standard; DNA; 50 BP.
XX
AC   AAQ69835;
XX
DT   25-MAR-2003  (revised)
DT   06-MAR-1995  (first entry)
XX
DE   Human papilloma virus type-16 E6/E7 (start site 97), target region.
XX
KW   DNA protein-binding assay; test sequence; screening sequence; promoter;
KW   target; TATA box; Herpes Simplex Virus; HSV; origin of replication; UL9;
KW   transcription factor; TFIID: ds.
XX
OS   Synthetic.
XX
PN   WO9414980-A1.
XX
PD   07-JUL-1994.
XX
PF   20-DEC-1993;   93WO-US012388.
XX
PR   23-DEC-1992;   92US-00996783.
PR   17-SEP-1993;   93US-00123936.
XX
PA   (GENE-) GENELABS TECHNOLOGIES INC.
XX
PI   Edwards CA,  Cantor CR,  Andrews BM,  Turin LM,  Fry KE;
XX
DR   WPI; 1994-234711/28.
XX
PT   Sequence-directed DNA-binding molecules - useful in pharmaceuticals and
PT   as molecular reagents.
XX
PS   Claim 28; Page 504; 587pp; English.
XX
CC   A DNA protein-binding assay is provided, useful for screening libraries
CC   of synthetic or biological cpds. for their ability to bind DNA test
CC   sequences. The assay is versatile in that any number of test sequences
CC   can be tested by placing the test sequence adjacent to a defined protein-
CC   binding screening sequence. Binding of mols. to these test sequences
CC   changes the binding characteristics of the protein mol. to its cognate
CC   binding sequence. When such a mol. binds the test sequence, the
CC   equilibrium of the DNA:protein complexes is disturbed, generating changes
CC   in the concentration of free DNA probe. One application of this method is
CC   to eucaryotic general transcription factors (e.g. TFIID), where the
CC   target region is typically selected from DNA sequences adjacent to the
CC   binding site for the eucaryotic transcription factor. Numerous exemplary
CC   test sequences are given: the sequences in AAQ69251-731 and AAQ69850
CC   correspond to promoter targets (typically, TATA box-contg. sites) for
CC   human genes and the sequences in AAQ69732-849 correspond to promoter
CC   targets for viral genes. The test sequences may also be randomly
CC   generated. DNA:protein interaction may be used for screening purposes,
CC   e.g. the Herpes Simplex Virus (HSV) origin of replication and UL9 (see
CC   AAQ69851-52, AAQ69865 and AAQ69891). (Updated on 25-MAR-2003 to correct
CC   PN field.)
XX
SQ   Sequence 50 BP; 22 A; 9 C; 10 G; 9 T; 0 U; 0 Other;

  Query Match             33.6%;  Score 14.8;  DB 2;  Length 50;
  Best Local Similarity   73.1%;  Pred. No. 1.9e+04;
  Matches   19;  Conservative    0;  Mismatches    7;  Indels    0;  Gaps    0;

Qy         11 TTCCTTCTTAATAACCGGTCGCGGTT 36
              | | ||  || ||||||||  |||||
Db         27 TGCTTTTATACTAACCGGTTTCGGTT 2


RESULT 24
AAT64297/c
ID   AAT64297 standard; DNA; 50 BP.
XX
AC   AAT64297;
XX
DT   25-MAR-2003  (revised)
DT   17-MAR-1997  (first entry)
XX
DE   HPV type 16 E6/E7 gene (start site 97) TFIID binding site.
XX
KW   Duplex DNA; target region; binding characteristic; DNA binding protein;
KW   TFIID; transcription factor; binding site; inhibition; enhance; cancer;
KW   inherited genetic disorder; ds.
XX
OS   Human papillomavirus.
XX
PN   US5578444-A.
XX
PD   26-NOV-1996.
XX
PF   20-DEC-1993;   93US-00171389.
XX
PR   27-JUN-1991;   91US-00723618.
PR   23-DEC-1992;   92US-00996783.
PR   17-SEP-1993;   93US-00123936.
XX
PA   (GENE-) GENELABS TECHNOLOGIES INC.
XX
PI   Fry KE,  Turin LM,  Andrews BM,  Cantor CR,  Edwards CA;
XX
DR   WPI; 1997-020402/02.
XX
PT   Altering binding characteristics of DNA binding proteins to duplex DNA -
PT   by attaching specific small cpd. to target region close to the protein's
PT   binding site, useful in treatment of viral disease, cancer etc.
XX
PS   Claim 6; Col 399; 264pp; English.
XX
CC   The sequences given in AAT63713-4312 represent duplex DNA's which act as
CC   target regions in the method of the invention. The method for altering
CC   the binding characteristics of a DNA-binding protein to duplex DNA
CC   comprises contacting the duplex DNA with a small molecule which binds
CC   sequence-specifically to a target region, where, when the small molecule
CC   is bound to the target region, it is adjacent to, but not overlapping by
CC   more than 4 bp, a binding site for a DNA-binding protein. The small
CC   molecule is added at a concentration effective to alter the binding of
CC   the DNA binding protein, pref. TFIID, to its binding site on the duplex
CC   DNA. The binding of the small molecule may inhibit or enhance the binding
CC   of the DNA-binding protein to its binding site. The compounds isolated
CC   using this method are potentially useful as therapeutic agents for
CC   treatment of any disease which involves a specific DNA sequence, e.g.
CC   cancer, or inherited genetic disorders etc. The method is suitable for
CC   screening large biological or chemical libraries and allows determination
CC   of sequence-specific and relative affinities of known DNA-binding agents
CC   for different DNA sequences. The design of these duplex DNA's allows a
CC   single DNA:protein interaction to be used for screening sequence-
CC   specific, or preferential, DNA binding proteins that recognise almost any
CC   possible sequence (see also AAT49539- 74). (Updated on 25-MAR-2003 to
CC   correct PF field.)
XX
SQ   Sequence 50 BP; 22 A; 9 C; 10 G; 9 T; 0 U; 0 Other;
```

```
  Query Match           33.6%;  Score 14.8;  DB 2;  Length 50;
  Best Local Similarity  73.1%;  Pred. No. 1.9e+04;
  Matches   19;  Conservative    0;  Mismatches    7;  Indels    0;  Gaps    0;

Qy         11 TTCCTTCTTAATAACCGGTCGCGGTT 36
              | | ||  || ||||||||  |||||
Db         27 TGCTTTTATACTAACCGGTTTCGGTT 2


RESULT 25
AAX17585/c
ID   AAX17585 standard; DNA; 50 BP.
XX
AC   AAX17585;
XX
DT   06-MAY-1999  (first entry)
XX
DE   Test sequence from human papilloma virus type-16 E6/E7 (start site 97).
XX
KW   Test sequence; DNA-binding molecule; screening sequence; human;
KW   nucleic acid amplification; target; viral; ds.
XX
OS   Human papillomavirus.
XX
PN   US5869241-A.
XX
PD   09-FEB-1999.
XX
PF   07-JUN-1995;   95US-00475228.
XX
PR   27-JUN-1991;   91US-00723618.
PR   23-DEC-1992;   92US-00996783.
PR   17-SEP-1993;   93US-00123936.
PR   20-DEC-1993;   93US-00171389.
XX
PA   (GENE-) GENELABS TECHNOLOGIES INC.
XX
PI   Fry KE,  Turin LM,  Andrews BM,  Cantor CR,  Edwards CA;
XX
DR   WPI; 1999-152755/13.
XX
PT   Determination of DNA sequence preference of a DNA-binding molecule -
PT   based on inhibition of binding of protein to oligonucleotide sequence
PT   attached to test sequence.
XX
PS   Claim 3; Col 399-400; 270pp; English.
XX
CC   Sequences AAX17001 to AAX17600 represent specifically claimed target test
CC   sequences that are used in the method of the invention of determining the
CC   DNA sequence preference of a DNA-binding molecule. The method comprises:
CC   (i) adding a test molecule and a DNA-binding protein to a mixture of
CC   duplex DNA test oligonucleotides, each of the test oligonucleotides
CC   having a test sequence adjacent to a screening sequence, where the
CC   screening sequence binds to the DNA-binding protein with a binding
CC   affinity that is independent of the DNA sequence of the test sequence,
CC   and where the mixture of duplex DNA test oligonucleotides includes
CC   several test sequences; (ii) incubating the test molecule, the mixture of
CC   duplex DNA test oligonucleotides and the DNA-binding protein for a time
CC   sufficient to permit binding of the test molecule to test sequences in
CC   the duplex DNA; (iii) separating unbound test oligonucleotides from test
CC   oligonucleotides bound to binding protein; (iv) amplifying the unbound
CC   test oligonucleotides; (v) repeating steps (ii) to (iv); (vi) isolating
CC   the amplified test oligonucleotides; and (vii) sequencing the isolated
CC   test oligonucleotides. Test sequences AAX17001-X17481 and AAX17600
CC   correspond to promoter targets for human genes and test sequences
CC   AAX17482-X17599 correspond to promoter targets for viral genes
XX
SQ   Sequence 50 BP; 22 A; 9 C; 10 G; 9 T; 0 U; 0 Other;

  Query Match           33.6%;  Score 14.8;  DB 2;  Length 50;
  Best Local Similarity  73.1%;  Pred. No. 1.9e+04;
  Matches   19;  Conservative    0;  Mismatches    7;  Indels    0;  Gaps    0;

Qy         11 TTCCTTCTTAATAACCGGTCGCGGTT 36
              | | ||  || ||||||||  |||||
Db         27 TGCTTTTATACTAACCGGTTTCGGTT 2


RESULT 26
ABK83076/c
ID   ABK83076 standard; DNA; 50 BP.
XX
AC   ABK83076;
XX
DT   27-AUG-2002  (first entry)
XX
DE   DNA binding molecule screening method test sequence #585.
XX
KW   DNA binding molecule screening; inhibition of transcription; infection;
KW   human immunodeficiency virus; HIV; parasite; cancer; cardiovascular;
KW   respiratory; gastrointestinal; endocrine; metabolic; rheumatic;
KW   immunological; haematological; neurological; psychiatric; dermatological;
KW   ophthalmological; musculo-skeletal; urogenital disorder; ss.
XX
OS   Synthetic.
XX
PN   US6384208-B1.
XX
PD   07-MAY-2002.
XX
PF   15-JUL-1999;   99US-00354947.
XX
PR   27-JUN-1991;   91US-00723618.
PR   23-DEC-1992;   92US-00996783.
PR   17-SEP-1993;   93US-00123936.
PR   20-DEC-1993;   93US-00171389.
PR   07-JUN-1995;   95US-00482080.
XX
PA   (GENE-) GENELABS TECHNOLOGIES INC.
XX
PI   Edwards CA,  Cantor CR,  Andrews BM,  Turin LM,  Fry KE;
XX
DR   WPI; 2002-442819/47.
XX
PT   Decreasing transcriptional activity of genes for treating infections or
PT   cancer, by administration of an agent that binds to two non-overlapping
PT   regions of the gene.
XX
PS   Example 15; SEQ ID NO 585; 98pp; English.
XX
CC   The invention relates to a method of decreasing transcriptional activity
CC   in a duplex deoxyribonucleic acid (DNA) template (T1) comprising
CC   contacting (T1) with a binding agent comprising at least one small duplex
CC   DNA-binding molecule (T2) coupled to at least one other small duplex-
CC   binding molecule that binds to a non-overlapping region of target
CC   sequence (TS). The method is useful for inhibiting transcription of a
CC   range of disease-related genes for treating infections (by viruses,
CC   including human immunodeficiency virus, bacteria, fungi, protozoa and
CC   parasites), cancer, cardiovascular, respiratory, gastrointestinal,
CC   endocrine/metabolic, rheumatic/immunological, haematological,
CC   neurological, psychiatric, dermatological, ophthalmological, musculo-
CC   skeletal, genetic or urogenital disorders. The method provides sequence-
CC   specific inhibition of transcription of pathological genes without
CC   affecting transcription of cellular genes regulated by the same
CC   transcription factor, and can be applied to regulation of any gene.
CC   ABK82492-ABK83155 represent DNA binding molecule test sequences used in
CC   the method of the invention
XX
SQ   Sequence 50 BP; 22 A; 9 C; 10 G; 9 T; 0 U; 0 Other;

  Query Match           33.6%;  Score 14.8;  DB 6;  Length 50;
  Best Local Similarity  73.1%;  Pred. No. 1.9e+04;
  Matches   19;  Conservative    0;  Mismatches    7;  Indels    0;  Gaps    0;

Qy         11 TTCCTTCTTAATAACCGGTCGCGGTT 36
              | | ||  || ||||||||  |||||
```

```
Db         27 TGCTTTTATACTAACCGGTTTCGGTT 2


RESULT 27
ABZ02818
ID   ABZ02818 standard; DNA; 50 BP.
XX
AC   ABZ02818;
XX
DT   09-JAN-2003  (first entry)
XX
DE   Human leukocyte gene expression profiling probe SEQ ID NO 2809.
XX
KW   T7; leukocyte; gene expression profiling; allograft rejection;
KW   atherosclerosis; congestive heart failure; systemic lupus erythematosus;
KW   rheumatoid arthritis; osteoarthritis; cytomegalovirus; infection; probe;
KW   ss.
XX
OS   Homo sapiens.
XX
PN   WO200257414-A2.
XX
PD   25-JUL-2002.
XX
PF   22-OCT-2001; 2001WO-US047856.
XX
PR   20-OCT-2000; 2000US-0241994P.
PR   08-JUN-2001; 2001US-0296764P.
XX
PA   (BIOC-) BIOCARDIA INC.
XX
PI   Wohlgemuth J,  Fry K,  Matcuk G,  Altman P,  Prentice J,  Phillips J;
PI   Ly N,  Woodward R,  Quertermous T,  Johnson F;
XX
DR   WPI; 2002-636525/68.
XX
PT   New system for leukocyte expression profiling, diagnosing a disease, or
PT   monitoring (the rate of) progression of a disease, e.g. atherosclerosis
PT   or congestive heart failure, comprises diagnostic oligonucleotides.
XX
PS   Claim 1; Page 417; 0pp; English.
XX
CC   The invention relates to a system for detecting gene expression, which
CC   comprises one or two isolated DNA molecules that detect expression of a
CC   gene, where the gene corresponds to any of 8143 oligonucleotides
CC   (ABZ00010-ABZ08152) each having 50 base pairs (bp). The system is useful
CC   for leukocyte expression profiling. It is particularly useful for
CC   diagnosing a disease, monitoring (rate of) progression of a disease,
CC   predicting therapeutic outcome, determining prognosis for a patient,
CC   predicting disease complications in an individual or monitoring response
CC   to treatment in an individual. The diseases include cardiac allograft
CC   rejection, kidney allograft rejection, liver allograft rejection,
CC   atherosclerosis, congestive heart failure, systemic lupus erythematosus,
CC   rheumatoid arthritis, osteoarthritis or cytomegalovirus infection
XX
SQ   Sequence 50 BP; 19 A; 6 C; 9 G; 16 T; 0 U; 0 Other;

  Query Match             33.6%;  Score 14.8;  DB 6;  Length 50;
  Best Local Similarity   64.7%;  Pred. No. 1.9e+04;
  Matches   22;  Conservative    0;  Mismatches   12;  Indels    0;  Gaps    0;

Qy         11 TTCCTTCTTAATAACCGGTCGCGGTTATTAAGAA 44
              || || |   || |  | | ||| ||| ||||||
Db         10 TTTCTGCAGCATTAAAGCTGGCGCTTAATAAGAA 43


RESULT 28
ADE80615/c
ID   ADE80615 standard; DNA; 50 BP.
XX
AC   ADE80615;
XX
DT   29-JAN-2004  (first entry)
XX
DE   Duplex oligonucleotide for DNA protein binding assay seq id 585.
XX
KW   DNA-binding; duplex DNA test oligonucleotide; DNA protein binding;
KW   library screening; promoter target; virus; ds.
XX
OS   Human papillomavirus.
XX
PN   US2003124530-A1.
XX
PD   03-JUL-2003.
XX
PF   13-NOV-2001; 2001US-00993346.
XX
PR   27-JUN-1991;   91US-00723618.
PR   23-DEC-1992;   92US-00996783.
PR   17-SEP-1993;   93US-00123936.
PR   20-DEC-1993;   93US-00171389.
PR   07-JUN-1995;   95US-00482080.
PR   15-JUL-1999;   99US-00354947.
XX
PA   (GENE-) GENELABS TECHNOLOGIES INC.
XX
PI   Edwards CA,  Cantor CR,  Andrews BM,  Turin LM,  Fry KE;
XX
DR   WPI; 2004-031270/03.
XX
PT   Screening for sequence-directed DNA-binding molecules comprises adding a
PT   test molecule to a test system composed of a DNA-binding protein and a
PT   duplex DNA test oligonucleotide having adjacent screening and test
PT   sequences.
XX
PS   Claim 2; SEQ ID NO 585; 283pp; English.
XX
CC   The invention describes a method for screening for molecules capable of
CC   binding to a selected test sequence in a duplex DNA. The above method
CC   comprises: constructing a duplex DNA test oligonucleotide having a
CC   screening sequence adjacent to a selected test sequence, where a DNA-
CC   binding protein is effective to bind to the screening sequence with a
CC   binding affinity that is substantially independent of such test sequence,
CC   but where DNA protein binding to the screening sequence is sensitive to
CC   binding of test molecules to such test sequence; adding a test molecule
CC   to be screened to a test system composed of the DNA-binding protein and
CC   the duplex DNA test oligonucleotide having the screening and test
CC   sequences adjacent one another; incubating the molecule in the test
CC   system for a period sufficient to permit binding of the molecule being
CC   tested to the test sequence in the duplex DNA; and comparing the amount
CC   of binding protein bound to the duplex DNA before and after the adding.
CC   The method is useful in screening libraries of synthetic or biological
CC   compounds for their ability to bind DNA test sequences. The method may
CC   also be used in characterising the preferred binding sequences of any
CC   selected DNA-binding molecule. This sequence represents a test sequence
CC   corresponding to a promoter target of a viral gene.
XX
SQ   Sequence 50 BP; 22 A; 9 C; 10 G; 9 T; 0 U; 0 Other;

  Query Match             33.6%;  Score 14.8;  DB 12;  Length 50;
  Best Local Similarity   73.1%;  Pred. No. 1.9e+04;
  Matches   19;  Conservative    0;  Mismatches    7;  Indels    0;  Gaps    0;

Qy         11 TTCCTTCTTAATAACCGGTCGCGGTT 36
              | | ||  || ||||||||  |||||
Db         27 TGCTTTTATACTAACCGGTTTCGGTT 2


RESULT 29
ADS91869/c
ID   ADS91869 standard; DNA; 50 BP.
XX
AC   ADS91869;
XX
DT   02-DEC-2004  (first entry)
XX
```

```
DE   Nematode LNA-modified oligo capture probe #212.
XX
KW   locked nucleic acid; LNA; biostability; cancer; cytostatic;
KW   diagnostic microarray; ss; probe; nematode.
XX
OS   Caenorhabditis elegans.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   modified_base   1
FT                   /*tag=  a
FT                   /mod_base= OTHER
FT                   /note= "OTHER= Locked nucleic acid (LNA). All LNA
FT                   cytosines are methyl cytosines."
FT   modified_base   2. .49
FT                   /*tag=  b
FT                   /mod_base= OTHER
FT                   /note= "OTHER= Locked nucleic acid (LNA) exists at every
FT                   third position (LNA-3). All LNA cytosines are methyl
FT                   cytosines."
XX
PN   WO2004035819-A2.
XX
PD   29-APR-2004.
XX
PF   21-OCT-2003; 2003WO-DK000715.
XX
PR   21-OCT-2002; 2002US-0420278P.
PR   19-MAY-2003; 2003DK-00000752.
XX
PA   (EXIQ-) EXIQON AS.
PA   (MORK/) MORK S.
XX
PI   Kauppinen S,  Alsbo C,  Nielsen PS,  Jeffares DC,  Mourier T;
PI   Arctander P,  Tommerup N,  Tolstrup N,  Vissing H;
XX
DR   WPI; 2004-357224/33.
XX
PT   New non-naturally-occurring nucleic acid having a melting temperature and
PT   capture efficiency higher than a control nucleic acid, useful for
PT   detecting and amplifying target nucleic acid, for alternative mRNA splice
PT   variant detection.
XX
PS   Example 8; Page 102; 400pp; English.
XX
CC   This invention relates to novel non-naturally occurring nucleic acids
CC   that exhibit enhanced biostability and capture efficiency i.e.
CC   hybridisation. Specifically, it refers to locked nucleic acid (LNA)
CC   oligos that have restricted flexibility in the ribofuranose ring of the
CC   nucleoside due to the presence of a 2'-O, 4'-C methylene bridge. The
CC   present invention describes these oligos as useful for detecting target
CC   nucleic acids and in particular for profiling mRNA splice variants,
CC   detecting mutations, deletions or duplications that may be associated
CC   with onset or increased risk of diseases such as cancer. Accordingly,
CC   such oligos exhibit cytostatic activities. Furthermore, these LNA
CC   containing oligos can be used as fluorescent in situ hybridisation (FISH)
CC   probes or on diagnostic microarrays to detect splice isoform signatures,
CC   as well as antisense oligos that can modulate or silence sense nucleic
CC   acid agents. This oligonucleotide sequence is a nematode LNA-modifed
CC   oligo capture probe of the invention.
XX
SQ   Sequence 50 BP; 13 A; 13 C; 13 G; 11 T; 0 U; 0 Other;

  Query Match             33.6%;  Score 14.8;  DB 13;  Length 50;
  Best Local Similarity   59.5%;  Pred. No. 1.9e+04;
  Matches   25;  Conservative    0;  Mismatches   17;  Indels    0;  Gaps    0;

Qy          1 GCGGGTCCCGTTCCTTCTTAATAACCGGTCGCGGTTATTAAG 42
              ||| | | |||||||   |  | |   |   |||||  | |||
Db         42 GCGTGCCACGTTCCTGTATCCTTAAAAGCGACGGTTTTGAAG 1


RESULT 30
ACI97514
ID   ACI97514 standard; DNA; 25 BP.
XX
AC   ACI97514;
XX
DT   14-OCT-2003  (first entry)
XX
DE   Human microarray DNA oligonucleotide SEQ ID NO 97505.
XX
KW   EST; ss; probe; expressed sequence tag; microarray; gene expression;
KW   genetic variation; biallelic marker; polymorphism; human;
KW   cross-species comparison.
XX
OS   Homo sapiens.
XX
PN   US2003104410-A1.
XX
PD   05-JUN-2003.
XX
PF   15-MAR-2002; 2002US-00098263.
XX
PR   16-MAR-2001; 2001US-0276759P.
XX
PA   (AFFY-) AFFYMETRIX INC.
XX
PI   Mittmann MP;
XX
DR   WPI; 2003-567953/53.
XX
PT   New array of nucleic acid probes, useful for in situ hybridization, in
PT   Southern, Northern or dot-blot hybridization to identify or detect the
PT   sequence or specific mutations of any gene.
XX
PS   Claim 1; SEQ ID NO 97505; 9pp; English.
XX
CC   The invention discloses a microarray comprising a plurality of nucleic
CC   acid probes including one of 2,018,500 fully defined sequences, or its
CC   perfect match, perfect mismatch, antisense match or antisense mismatch.
CC   Also disclosed is a method of gene expression analysis. The array is used
CC   in monitoring gene expression levels by hybridisation to a DNA library,
CC   in analysis of genetic variation or in hybridisation of tag-labelled
CC   compounds. The nucleic acid probes are specifically designed for analysis
CC   of at least one target sequence. The method of analysis comprises
CC   hybridising at least one or more nucleic acids to at least two or more
CC   nucleic acid probes and detecting the hybridisation. The nucleic acid
CC   probes are attached to a solid support. The analysis comprises monitoring
CC   gene expression levels, identifying biallelic markers or polymorphisms,
CC   or family members of a gene and a cross-species comparison. Each of the
CC   nucleic acids further comprises a tag sequence. The array of nucleic acid
CC   probes is useful in in situ hybridisation, in Southern, Northern or dot-
CC   blot hybridisation to identify or detect the sequence or specific
CC   mutations of any gene, in mapping the 5' termini of mRNA molecules by
CC   primer extensions or in screening cDNA or genomic libraries or subclones
CC   for additional subclones containing segments of DNA that have been
CC   isolated and previously sequenced. The sequence presented is one of the
CC   nucleic acid probes incorporated in the microarray. Note: The sequence
CC   data for this patent can also be obtained in electronic format directly
CC   from USPTO at seqdata.uspto.goc/sequence.html
XX
SQ   Sequence 25 BP; 2 A; 9 C; 5 G; 9 T; 0 U; 0 Other;

  Query Match             33.2%;  Score 14.6;  DB 9;  Length 25;
  Best Local Similarity   81.0%;  Pred. No. 2.1e+04;
  Matches   17;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          2 CGGGTCCCGTTCCTTCTTAAT 22
              |||||||||||| ||  || |
Db          3 CGGGTCCCGTTCGTTACTACT 23


RESULT 31
AAQ10823/c
ID   AAQ10823 standard; DNA; 34 BP.
```

XX
AC AAQ10823;
XX
DT 09-MAY-1991 (first entry)
XX
DE Pneumocystis carinii 18S rRNA-targetted probe, 1485.
XX
KW Hybridisation assay; pneumonia; AIDS; ss.
XX
OS Synthetic.
XX
PN WO9102092-A.
XX
PD 21-FEB-1991.
XX
PF 11-AUG-1989; 89US-00392679.
XX
PR 11-AUG-1989; 89US-00392679.
XX
PA (GENE-) GENE-TRAK SYST.
XX
PI Shah JS, Buharin A, Lane DJ;
XX
DR WPI; 1991-073563/10.
XX
PT Nucleic acid fragment - capable of hybridising to r RNA or r DNA of
PT pneumocystis carinii, useful as probes for detection of P carinii.
XX
PS Disclosure; Page 11; 36pp; English.
XX
CC This oligonucleotide has a sequence specific for a region of the rDNA of
CC Pneumocystis carinii, the causative agent of pneumonia. It can be used as
CC a probe in hybridisation assays to detect P. carinii in clinical samples.
CC This probe detects human but not most non-human P.carinii isolates. See
CC also AAQ10820-22 and AAQ10824-33
XX
SQ Sequence 34 BP; 12 A; 7 C; 11 G; 4 T; 0 U; 0 Other;

Query Match 33.2%; Score 14.6; DB 2; Length 34;
Best Local Similarity 81.0%; Pred. No. 2.2e+04;
Matches 17; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 9 CGTTCCTTCTTAATAACCGGT 29
| ||||||||| || ||||||
Db 31 CCTTCCTTCTGGATTACCGGT 11

RESULT 32
AAT42447/c
ID AAT42447 standard; DNA; 34 BP.
XX
AC AAT42447;
XX
DT 25-MAR-2003 (revised)
DT 08-JAN-1997 (first entry)
XX
DE Probe 1485 for P. carinii 18S rRNA.
XX
KW Probe; pneumocystis carinii; 18S rRNA; human; mammal; immunodeficiency;
KW pneumonia; ferret; rat; ss.
XX
OS Synthetic.
XX
PN US5519127-A.
XX
PD 21-MAY-1996.
XX
PF 21-JAN-1992; 92US-00826657.
XX
PR 11-AUG-1989; 89US-00392679.
XX
PA (STAD) AMOCO CORP.
XX
PI Lane DJ, Buharin A, Shah J;
XX
DR WPI; 1996-259122/26.
XX
PT Nucleic acid probes specific for human Pneumocystis carinii - provide
PT sensitive, accurate and rapid diagnosis of infection.
XX
PS Claim 1; Col 13-14; 15pp; English.
XX
CC AAT42443-T42456 represent probes for human Pneumocystis carinii (Pc) 18S
CC rRNA. Pc infects humans and most mammalian hosts, but rarely cause
CC illness in normal individuals. However, in certain conditions of
CC immunodeficiency, Pc does give rise to life threatening pneumonia. This
CC sequence hybridises to the 18S rRNA at positions 641-652, and is capable
CC of hybridisation to human Pc. These probes can be used as a probe set for
CC a Pc assay. The 1485 (this sequence), 1487 (see AAT42452) and 1159 (see
CC AAT42448) probes are reactive with all human Pc isolates. The probes 1485
CC and 1487 hybridise to rRNA and rDNA of human Pc, but not to other fungi
CC or bacteria. The 1485 and 1487 probes can therefore be used to detect Pc
CC in clinical samples. The rest of these probe sequences are reactive
CC mainly with ferret Pc and non-human strains of Pc. By using these probe
CC sequences, a more sensitive, accurate and rapid diagnosis can be
CC performed, with reduced expense, in comparison to current technology.
CC rRNA is present in the cell at high concentration, and is not likely to
CC undergo lateral transfer. (Updated on 25-MAR-2003 to correct PF field.)
XX
SQ Sequence 34 BP; 12 A; 7 C; 11 G; 4 T; 0 U; 0 Other;

Query Match 33.2%; Score 14.6; DB 2; Length 34;
Best Local Similarity 81.0%; Pred. No. 2.2e+04;
Matches 17; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 9 CGTTCCTTCTTAATAACCGGT 29
| ||||||||| || ||||||
Db 31 CCTTCCTTCTGGATTACCGGT 11

RESULT 33
ADI35720/c
ID ADI35720 standard; DNA; 35 BP.
XX
AC ADI35720;
XX
DT 15-APR-2004 (first entry)
XX
DE Human tyrosine kinase receptor TrkBIg26His reverse PCR primer.
XX
KW tyrosine kinase receptor; biosensor; human; PCR; primer; ss.
XX
OS Homo sapiens.
OS Synthetic.
XX
PN WO2003025016-A2.
XX
PD 27-MAR-2003.
XX
PF 17-SEP-2002; 2002WO-GB004214.
XX
PR 17-SEP-2001; 2001GB-00022400.
XX
PA (UYBR-) UNIV BRISTOL.
XX
PI Dawbarn D, Allen SJ, Robertson AGS;
XX
DR WPI; 2003-371797/35.
XX
PT Producing tyrosine kinase receptor-related polypeptides useful as
PT biosensors, comprises expressing a tyrosine kinase receptor-related
PT polypeptide in a recombinant expression system.
XX
PS Disclosure; Page 15; 54pp; English.
XX
CC The present invention describes a method for producing tyrosine kinase

```
CC   receptor-related polypeptides comprising expressing a tyrosine kinase
CC   receptor-related polypeptide in a recombinant expression system,
CC   separating the expressed monomeric tyrosine kinase receptor-related
CC   polypeptide from multimeric form(s) of the expressed polypeptide, and
CC   allowing refolding of the expressed tyrosine kinase receptor-related
CC   polypeptide into a biologically active form. Also described: (1)
CC   purifying recombinant TrkAIg2, TrkAIg2.6, TrkBIga, or TrkCIga from
CC   inclusion bodies in a bacterial expression system in which monomeric
CC   TrkAIg2, TrkAIg2.6, TrkBIga, or TrkCIga is separated from a mixture
CC   including monomeric and multimeric TrkAIg2, TrkAIg2.6, TrkBIga, or
CC   TrkCIga by a gel filtration step, and allowed to refold into an active
CC   form; (2) preparations of TrkAIg2, TrkAIg2.6, TrkBIg2 or TrkCIg2 obtained
CC   by a method above comprising less than 20% TrkAIg2, TrkAIg2.6, TrkBIg2 or
CC   TrkCIg2 dimer or dimer aggregate; and (3) preparations of TrkAIg2,
CC   TrkAIg2.6, TrkBIg2 or TrkCIg2 obtained by a method above comprising more
CC   than 80-99% TrkAIg2, TrkAIg2.6, TrkBIg2 or TrkCIg2, or comprising 100%
CC   TrkAIg2, TrkAIg2.6, TrkBIg2 or TrkCIg2 monomer. The method is useful for
CC   producing tyrosine kinase receptor-related polypeptides which can be used
CC   as biosensors. Compared with previous methods, the present method
CC   provides improved yields, has improved stability, does not require a
CC   separate dialysis-based refolding step, is much quicker than processes
CC   involving dialysis, produces product at higher concentrations without
CC   strand swap dimers, and since product is not in contact with urea for the
CC   length periods required during a dialysis procedure, it is less likely to
CC   be amidated. The present sequence represents a PCR primer for
CC   TrkBIg26His, which is used in the exemplification of the present
CC   invention.
XX
SQ   Sequence 35 BP; 6 A; 10 C; 7 G; 10 T; 0 U; 2 Other;

  Query Match             33.2%;  Score 14.6;  DB 10;  Length 35;
  Best Local Similarity   73.9%;  Pred. No. 2.2e+04;
  Matches   17;  Conservative    1;  Mismatches    5;  Indels    0;  Gaps    0;

Qy        21 ATAACCGGTCGCGGTTATTAAGA 43
             | |||||||||| | : ||||| |
Db        32 AAAACCGGTCGGGYYCATTAATA 10


RESULT 34
AAV06323/c
ID   AAV06323 standard; DNA; 38 BP.
XX
AC   AAV06323;
XX
DT   06-MAY-1998  (first entry)
XX
DE   Human Col III gene 3' end synthesising 3' primer.
XX
KW   Collagen; Col III; recombinant; post-translational enzyme; human;
KW   procollagen; PCR primer; ss.
XX
OS   Synthetic.
OS   Homo sapiens.
XX
PN   WO9738710-A1.
XX
PD   23-OCT-1997.
XX
PF   11-APR-1997;   97WO-US007300.
XX
PR   12-APR-1996;   96US-00631336.
XX
PA   (FIBR-) FIBROGEN INC.
PA   (FIFI-) ACAD FINLAND.
XX
PI   Kivirikki KI,  Pihlajaniemi T;
XX
DR   WPI; 1997-526203/48.
XX
PT   Recombinant production of (pro)collagen having correct folding - using
PT   vectors encoding collagen sub:unit and collagen post-translational enzyme
PT   respectively.
XX
PS   Example 11; Page 67; 90pp; English.
XX
CC   This primer is used to synthesise the 3' end of the human Col III gene by
CC   PCR amplification. This is used in the construction of recombinant
CC   vectors containing collagen genes. A novel method for producing a
CC   (pro)collagen polypeptide comprises culturing a host cell, where the host
CC   cell has been infected, transfected or transformed with a first
CC   expression vector comprising a polynucleotide molecule having a nucleic
CC   acid sequence which encodes a (pro)collagen subunit and a second
CC   expression vector comprising a polynucleotide molecule having a nucleic
CC   acid sequence which encodes at least one (pro)collagen post-translational
CC   enzyme or enzyme subunit. The (pro)collagen polypeptide is then purified
CC   from the cultured cell. The (pro)collagen polypeptide is selected from
CC   collagen types IV, V, VI, VII, VIII, IX, X, XI, XII, XIII, XIV, XV, XVI,
CC   XVII, XVIII, and XIX. The methods can be used for the production of
CC   collagens such as human collagens which can be used in therapeutic
CC   applications. The method provides for the synthesis of correctly folded
CC   proteins so that they exhibit the normal triple-helical conformation
CC   characteristic of procollagens and collagens. Purification of the
CC   collagens is greatly facilitated
XX
SQ   Sequence 38 BP; 12 A; 10 C; 11 G; 5 T; 0 U; 0 Other;

  Query Match             33.2%;  Score 14.6;  DB 2;  Length 38;
  Best Local Similarity   69.0%;  Pred. No. 2.2e+04;
  Matches   20;  Conservative    0;  Mismatches    9;  Indels    0;  Gaps    0;

Qy         4 GGTCCCGTTCCTTCTTAATAACCGGTCGC 32
             || || |||   | || |||| ||| |||
Db        38 GGCCCTGTTTGCTTTTTATAAGCGGCCGC 10


RESULT 35
ABZ47662
ID   ABZ47662 standard; DNA; 41 BP.
XX
AC   ABZ47662;
XX
DT   26-JUN-2003  (first entry)
XX
DE   Human ATP-binding cassette ABCC7/CFTR gene polymorphic site, #4446.
XX
KW   Human; drug metabolising enzyme; gene; drug metabolism; chromosome 7;
KW   polymorphic site; drug evaluation; drug screening; genotyping;
KW   genetic profiling; therapeutic customisation; adverse reaction;
KW   clinical trial; drug approval; single nucleotide polymorphism; SNP; ds.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   variation       replace(21,G)
FT                   /*tag=  a
FT                   /standard_name= "Single nucleotide polymorphism (SNP)"
XX
PN   WO200252044-A2.
XX
PD   04-JUL-2002.
XX
PF   27-DEC-2001; 2001WO-JP011592.
XX
PR   27-DEC-2000; 2000JP-00399443.
PR   02-MAY-2001; 2001JP-00135256.
PR   27-AUG-2001; 2001JP-00256862.
XX
PA   (RIKE ) RIKEN KK.
XX
PI   Nakamura Y,  Sekine A,  Iida A,  Saito S;
XX
DR   WPI; 2002-583571/62.
XX
PT   Identifying individuals having a polymorphism, useful for determining the
PT   effectiveness or side effect of a drug or treatment protocol, comprises
```

PT detecting at least one polymorphism in the drug metabolizing enzyme
PT nucleic acid.
XX
PS Claim 23; Page 147; 2785pp; English.
XX
CC Sequences ABZ43217-ABZ50887 represent polymorphic sites within genes
CC encoding enzymes associated with drug metabolism. The invention relates
CC to methods and compositions for identifying individuals who have at least
CC one polymorphism in such drug metabolising enzyme-encoding genes. The
CC polymorphisms may be identified in a nucleic acid sample using probes or
CC primers specific for a sequence selected from ABZ43217-ABZ50887 using a
CC variety of detection assays, including hybridisation assays, nucleic acid
CC arrays and PCR-based methods. The invention also encompasses methods of
CC evaluating and screening drugs using genetic polymorphism data. Genetic
CC polymorphism data, particularly that relating to single nucleotide
CC polymorphisms (SNPs), may be used in studying the relationship between
CC DNA sequence variations and human diseases, conditions, and responses to
CC drugs. SNPs are also useful as polymorphism markers for discovering genes
CC that cause or exacerbate certain diseases. SNPs are particularly useful
CC in the above respects as they are stable in populations, occur
CC frequently, and have lower mutation rates than other genome variations
CC such as repeating sequences. The detection and analysis of polymorphisms
CC in genes encoding drug metabolising enzymes allows the customisation of
CC drug therapies based upon the genetic profile of individual patients.
CC This would not only take the guesswork out of selecting the drug with the
CC greatest therapeutic effect for a particular patient, but would also
CC reduce the likelihood of adverse reactions, thereby increasing safety.
CC Methods of the invention are also useful in the drug discovery and
CC approval processes. For example, individuals could be selected for
CC clinical trials only if their genetic profiles indicate that they are
CC capable of responding to a particular drug or drug class, and previously
CC failed drug candidates could be revived if they were matched with more
CC appropriate patient populations. The methods, data and compositions of
CC the invention may therefore lead to a an increase in the range of
CC possible drug targets and decreases in the number of adverse drug
CC reactions, failed drug trials, the time taken for a drug to be approved,
CC the length of time patients are on medication and the number of different
CC medications a patient needs to take before finding an effective therapy
XX
SQ Sequence 41 BP; 8 A; 15 C; 2 G; 16 T; 0 U; 0 Other;

Query Match 33.2%; Score 14.6; DB 6; Length 41;
Best Local Similarity 81.0%; Pred. No. 2.3e+04;
Matches 17; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 5 GTCCCGTTCCTTCTTAATAAC 25
||||| ||||||| | ||| |
Db 16 GTCCCATTCCTTCATTATATC 36

RESULT 36
ABZ45067
ID ABZ45067 standard; DNA; 41 BP.
XX
AC ABZ45067;
XX
DT 26-JUN-2003 (first entry)
XX
DE Human ATP-binding cassette ABCC7/CFTR gene polymorphic site, #1851.
XX
KW Human; drug metabolising enzyme; gene; drug metabolism; chromosome 7;
KW polymorphic site; drug evaluation; drug screening; genotyping;
KW genetic profiling; therapeutic customisation; adverse reaction;
KW clinical trial; drug approval; single nucleotide polymorphism; SNP; ds.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT variation replace(21,G)
FT /*tag= a
FT /standard_name= "Single nucleotide polymorphism (SNP)"
XX
PN WO200252044-A2.
XX
PD 04-JUL-2002.
XX
PF 27-DEC-2001; 2001WO-JP011592.
XX
PR 27-DEC-2000; 2000JP-00399443.
PR 02-MAY-2001; 2001JP-00135256.
PR 27-AUG-2001; 2001JP-00256862.
XX
PA (RIKE) RIKEN KK.
XX
PI Nakamura Y, Sekine A, Iida A, Saito S;
XX
DR WPI; 2002-583571/62.
XX
PT Identifying individuals having a polymorphism, useful for determining the
PT effectiveness or side effect of a drug or treatment protocol, comprises
PT detecting at least one polymorphism in the drug metabolizing enzyme
PT nucleic acid.
XX
PS Claim 23; Page 94; 2785pp; English.
XX
CC Sequences ABZ43217-ABZ50887 represent polymorphic sites within genes
CC encoding enzymes associated with drug metabolism. The invention relates
CC to methods and compositions for identifying individuals who have at least
CC one polymorphism in such drug metabolising enzyme-encoding genes. The
CC polymorphisms may be identified in a nucleic acid sample using probes or
CC primers specific for a sequence selected from ABZ43217-ABZ50887 using a
CC variety of detection assays, including hybridisation assays, nucleic acid
CC arrays and PCR-based methods. The invention also encompasses methods of
CC evaluating and screening drugs using genetic polymorphism data. Genetic
CC polymorphism data, particularly that relating to single nucleotide
CC polymorphisms (SNPs), may be used in studying the relationship between
CC DNA sequence variations and human diseases, conditions, and responses to
CC drugs. SNPs are also useful as polymorphism markers for discovering genes
CC that cause or exacerbate certain diseases. SNPs are particularly useful
CC in the above respects as they are stable in populations, occur
CC frequently, and have lower mutation rates than other genome variations
CC such as repeating sequences. The detection and analysis of polymorphisms
CC in genes encoding drug metabolising enzymes allows the customisation of
CC drug therapies based upon the genetic profile of individual patients.
CC This would not only take the guesswork out of selecting the drug with the
CC greatest therapeutic effect for a particular patient, but would also
CC reduce the likelihood of adverse reactions, thereby increasing safety.
CC Methods of the invention are also useful in the drug discovery and
CC approval processes. For example, individuals could be selected for
CC clinical trials only if their genetic profiles indicate that they are
CC capable of responding to a particular drug or drug class, and previously
CC failed drug candidates could be revived if they were matched with more
CC appropriate patient populations. The methods, data and compositions of
CC the invention may therefore lead to a an increase in the range of
CC possible drug targets and decreases in the number of adverse drug
CC reactions, failed drug trials, the time taken for a drug to be approved,
CC the length of time patients are on medication and the number of different
CC medications a patient needs to take before finding an effective therapy
XX
SQ Sequence 41 BP; 8 A; 15 C; 2 G; 16 T; 0 U; 0 Other;

Query Match 33.2%; Score 14.6; DB 6; Length 41;
Best Local Similarity 81.0%; Pred. No. 2.3e+04;
Matches 17; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 5 GTCCCGTTCCTTCTTAATAAC 25
||||| ||||||| | ||| |
Db 16 GTCCCATTCCTTCATTATATC 36

RESULT 37
ADL60971/c
ID ADL60971 standard; DNA; 41 BP.
XX
AC ADL60971;
XX

DT 01-JUL-2004 (first entry)
XX
DE Human organic anion transport protein SNP region #680.
XX
KW gene therapy; human; OATP2; cMOAT; hepatic disease; metabolic disease;
KW inflammatory disease; cardiovascular disease; hyperproliferative disease;
KW neurological disease; infectious disease; liver disease;
KW high cholesterol; hypertension; congestive heart failure;
KW coronary heart disease; cancer; wound healing; ds; SNP;
KW single nucleotide polymorphism.
XX
OS Homo sapiens.
XX
PN US2004068096-A1.
XX
PD 08-APR-2004.
XX
PF 20-SEP-2002; 2002US-00252155.
XX
PR 21-SEP-2001; 2001US-0324172P.
PR 27-NOV-2001; 2001US-0333700P.
XX
PA (TSUC/) TSUCHIHASHI Z.
PA (HUIL/) HUI L.
PA (KIRC/) KIRCHGESSNER T.
XX
PI Tsuchihashi Z, Hui L, Kirchgessner T;
XX
DR WPI; 2004-304621/28.
XX
PT New nucleic acid encoding human OATP2 or cMOAT protein, useful in
PT diagnosing, treating or preventing diseases or disorders, e.g.
PT inflammatory, cardiovascular, hyperproliferative, neurological or
PT infectious diseases.
XX
PS Disclosure; SEQ ID NO 747; 296pp; English.
XX
CC The invention relates to an isolated nucleic acid derived from a human
CC gene encoding a protein, i.e. human OATP2 protein or human cMOAT protein,
CC where the nucleic acid comprises at least one polymorphic position The
CC nucleic acid and the encoded protein, kits and composition are useful in
CC diagnosing, treating or preventing diseases or disorders, e.g. hepatic,
CC metabolic, inflammatory, cardiovascular, hyperproliferative,
CC neurological, infectious diseases, liver disease, high cholesterol,
CC hypertension, congestive heart failure or coronary heart disease and
CC cancer and promotes wound healing. The present sequence represents the
CC amino acid sequence of a human organic anion transport protein.
XX
SQ Sequence 41 BP; 16 A; 6 C; 14 G; 5 T; 0 U; 0 Other;

Query Match 33.2%; Score 14.6; DB 12; Length 41;
Best Local Similarity 69.0%; Pred. No. 2.3e+04;
Matches 20; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

Qy 10 GTTCCTTCTTAATAACCGGTCGCGGTTAT 38
| |||| ||| ||||| | |||| ||
Db 41 GCTCCTCCTTTTTAACCTCTACCGGTCAT 13

RESULT 38
ABN84317/c
ID ABN84317 standard; DNA; 45 BP.
XX
AC ABN84317;
XX
DT 29-AUG-2003 (revised)
DT 01-OCT-2002 (first entry)
XX
DE Rhinovirus specific PCR primer rh1D.R1.
XX
KW cW985; perturbagen; viral infection; virucide; human; gene therapy; PCR;
KW primer; ss.
XX
OS Human rhinovirus sp.
XX
PN WO200255697-A2.
XX
PD 18-JUL-2002.
XX
PF 16-NOV-2001; 2001WO-US043486.
XX
PR 27-NOV-2000; 2000US-0253333P.
PR 28-FEB-2001; 2001US-0272026P.
XX
PA (DELT-) DELTAGEN PROTEOMICS INC.
XX
PI Kamb CA, Poritz MA, Teng DH;
XX
DR WPI; 2002-557822/59.
XX
PT New cW985 perturbagen polypeptides and polynucleotides useful for
PT treating viral infections by the picornaviridae class, in chromosomal
PT mapping, tissue typing, forensic biology, or viral serotyping.
XX
PS Example 6; Fig 17; 117pp; English.
XX
CC The present sequence is of PCR primer rh1D.R1, which was used in an
CC example from the invention describing methods for identifying the targets
CC of viral-neutralising perturbagens. In order to construct a viral target
CC library, 10 of the polypeptides encoded by the human rhinovirus-14 genome
CC were RT-PCR amplified from viral RNA using the viral-specific
CC oligonucleotides given in ABN84310-29. These include the present primer
CC which, in addition to the viral sequence, includes a 5' restriction site.
CC PCR products were cloned into pVT725 (HIS+) in-frame with the LexA
CC binding domain. The polynucleotide sequence (see ABN84303) encoding novel
CC perturbagen cW985 (see ABB79541) was cloned into pVT578 such that the 53
CC amino acids of cW985 were fused in-frame with the C-terminus of the LexA
CC activating domain. The pVT578-W985 and pVT725-viral library constructs
CC were introduced into a yeast strain, and viral proteins that interacted
CC with the W985 perturbagen were identified. The invention provides host
CC cells, vectors and gene therapy vectors comprising polynucleotides
CC encoding cW985. The host cells provide for methods for producing
CC polypeptides having viral-related activity, which in turn can be used to
CC identify potential therapeutics. The invention also provides methods for
CC identifying a cellular target that interacts with the perturbagen, and
CC for using such targets to screen for viral therapeutics. (Updated on 29-
CC AUG-2003 to standardise OS field)
XX
SQ Sequence 45 BP; 11 A; 12 C; 7 G; 15 T; 0 U; 0 Other;

Query Match 33.2%; Score 14.6; DB 6; Length 45;
Best Local Similarity 62.2%; Pred. No. 2.3e+04;
Matches 23; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

Qy 4 GGTCCCGTTCCTTCTTAATAACCGGTCGCGGTTATTA 40
||| | || || || ||| ||| ||| || | |
Db 37 GGTGACATTAAATCCTATTAAGCGGCCGCTGTATTGA 1

RESULT 39
ADO18016
ID ADO18016 standard; DNA; 47 BP.
XX
AC ADO18016;
XX
DT 01-JUL-2004 (first entry)
XX
DE Primer of the invention #242.
XX
KW single nucleotide polymorphism; primer; ss.
XX
OS Synthetic.
XX
PN WO2004003220-A2.
XX
PD 08-JAN-2004.

XX
PF 26-JUN-2003; 2003WO-US020150.
XX
PR 28-JUN-2002; 2002US-0392504P.
XX
PA (ORCH-) ORCHID BIOSCIENCES INC.
XX
PI Giles R, Baisch JM, Mckeown B, Stolorow M;
XX
DR WPI; 2004-091088/09.
XX
PT New panel of single nucleotide polymorphisms comprising two or more
PT single nucleotide polymorphisms, useful for analyzing compromised nucleic
PT acid samples.
XX
PS Claim 2; SEQ ID NO 243; 76pp; English.
XX
CC The present invention relates to a panel of two or more single nucleotide
CC polymorphisms, where each of the polymorphisms of the panel are selected
CC from single nucleotide polymorphisms that are not genetically linked with
CC respect to one another, and where each of the polymorphisms of the panel
CC are selected from single nucleotide polymorphisms that are located
CC outside tandem repeat nucleic acid sequences. The known sample and the
CC unknown sample are from the same individual. The known sample is from a
CC family member. The compromised nucleic acid sample comprises nucleic acid
CC fragments from 10-100 nucleotides in length. The identity of the one or
CC more single nucleotide polymorphisms is determined using a single base
CC primer extension reaction. The present sequence represents a primer of
CC the invention.
XX
SQ Sequence 47 BP; 6 A; 12 C; 17 G; 11 T; 0 U; 1 Other;

Query Match 33.2%; Score 14.6; DB 12; Length 47;
Best Local Similarity 66.7%; Pred. No. 2.3e+04;
Matches 20; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Qy 3 GGGTCCCGTTCCTTCTTAATAACCGGTCGC 32
|| | |||||| || || || | || | |
Db 1 GGATGGCGTTCCGTCCTATTATCTGGGCNC 30

RESULT 40
ADO18088
ID ADO18088 standard; DNA; 47 BP.
XX
AC ADO18088;
XX
DT 01-JUL-2004 (first entry)
XX
DE Primer of the invention #314.
XX
KW single nucleotide polymorphism; primer; ss.
XX
OS Synthetic.
XX
PN WO2004003220-A2.
XX
PD 08-JAN-2004.
XX
PF 26-JUN-2003; 2003WO-US020150.
XX
PR 28-JUN-2002; 2002US-0392504P.
XX
PA (ORCH-) ORCHID BIOSCIENCES INC.
XX
PI Giles R, Baisch JM, Mckeown B, Stolorow M;
XX
DR WPI; 2004-091088/09.
XX
PT New panel of single nucleotide polymorphisms comprising two or more
PT single nucleotide polymorphisms, useful for analyzing compromised nucleic
PT acid samples.
XX
PS Claim 2; SEQ ID NO 315; 76pp; English.
XX
CC The present invention relates to a panel of two or more single nucleotide
CC polymorphisms, where each of the polymorphisms of the panel are selected
CC from single nucleotide polymorphisms that are not genetically linked with
CC respect to one another, and where each of the polymorphisms of the panel
CC are selected from single nucleotide polymorphisms that are located
CC outside tandem repeat nucleic acid sequences. The known sample and the
CC unknown sample are from the same individual. The known sample is from a
CC family member. The compromised nucleic acid sample comprises nucleic acid
CC fragments from 10-100 nucleotides in length. The identity of the one or
CC more single nucleotide polymorphisms is determined using a single base
CC primer extension reaction. The present sequence represents a primer of
CC the invention.
XX
SQ Sequence 47 BP; 6 A; 12 C; 17 G; 11 T; 0 U; 1 Other;

Query Match 33.2%; Score 14.6; DB 12; Length 47;
Best Local Similarity 66.7%; Pred. No. 2.3e+04;
Matches 20; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Qy 3 GGGTCCCGTTCCTTCTTAATAACCGGTCGC 32
|| | |||||| || || || | || | |
Db 1 GGATGGCGTTCCGTCCTATTATCTGGGCNC 30

Search completed: May 24, 2005, 12:14:32
Job time : 271 secs

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: May 24, 2005, 11:49:09 ; Search time 98 Seconds
(without alignments)
734.655 Million cell updates/sec

Title: US-10-673-063-3_COPY_900_943
Perfect score: 44
Sequence: 1 gcgggtcccgttccttctta.....ccggtcgcggttattaagaa 44

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 1209694

Minimum DB seq length: 0
Maximum DB seq length: 50

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 100 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
c	1	15.8	35.9	25	4	US-09-396-196G-29372	Sequence 29372, A
	2	15.2	34.5	33	6	5519127-36	Patent No. 5519127
	3	15.2	34.5	33	6	5519127-36	Patent No. 5519127
c	4	14.8	33.6	25	4	US-09-396-196G-29373	Sequence 29373, A
	5	14.8	33.6	36	4	US-09-665-189A-40	Sequence 40, Appl
c	6	14.8	33.6	47	4	US-09-422-978-240	Sequence 240, App
c	7	14.8	33.6	47	4	US-09-422-978-1725	Sequence 1725, Ap
c	8	14.8	33.6	50	1	US-08-171-389-585	Sequence 585, App
c	9	14.8	33.6	50	1	US-08-123-936-585	Sequence 585, App
c	10	14.8	33.6	50	2	US-08-475-228A-585	Sequence 585, App
c	11	14.8	33.6	50	3	US-08-482-080A-585	Sequence 585, App
c	12	14.8	33.6	50	3	US-09-354-947-585	Sequence 585, App
c	13	14.8	33.6	50	5	PCT-US93-12388-585	Sequence 585, App
	14	14.6	33.2	33	6	5519127-17	Patent No. 5519127
	15	14.6	33.2	33	6	5519127-17	Patent No. 5519127
c	16	14.6	33.2	34	6	5519127-4	Patent No. 5519127
c	17	14.6	33.2	34	6	5519127-4	Patent No. 5519127
	18	14.6	33.2	39	6	5519127-33	Patent No. 5519127
	19	14.6	33.2	39	6	5519127-33	Patent No. 5519127
	20	14.6	33.2	46	1	US-07-826-928A-30	Sequence 30, Appl
	21	14.6	33.2	47	4	US-09-422-978-3415	Sequence 3415, Ap
	22	14.6	33.2	49	3	US-08-889-502-34	Sequence 34, Appl
	23	14.4	32.7	47	4	US-09-422-978-2970	Sequence 2970, Ap
c	24	14.2	32.3	20	4	US-09-198-452A-6719	Sequence 6719, Ap
c	25	14.2	32.3	20	4	US-09-922-146-31	Sequence 31, Appl
c	26	14.2	32.3	21	4	US-09-657-472-1129	Sequence 1129, Ap
	27	14.2	32.3	36	1	US-07-956-697B-3	Sequence 3, Appli
	28	14.2	32.3	36	1	US-08-263-098-3	Sequence 3, Appli
c	29	14.2	32.3	38	2	US-08-857-946-83	Sequence 83, Appl
c	30	14.2	32.3	38	3	US-08-970-740-83	Sequence 83, Appl
c	31	14.2	32.3	49	3	US-09-538-709-983	Sequence 983, App
	32	14	31.8	24	3	US-08-379-452-17	Sequence 17, Appl
	33	14	31.8	24	3	US-09-409-670-17	Sequence 17, Appl
	34	14	31.8	30	3	US-09-461-697-456	Sequence 456, App
	35	14	31.8	32	4	US-09-709-103-34	Sequence 34, Appl
	36	14	31.8	32	4	US-09-439-410A-34	Sequence 34, Appl
	37	14	31.8	35	4	US-09-709-103-8	Sequence 8, Appli
	38	14	31.8	35	4	US-09-439-410A-8	Sequence 8, Appli
c	39	14	31.8	41	4	US-09-060-299-117	Sequence 117, App
c	40	14	31.8	41	4	US-09-402-923A-117	Sequence 117, App
	41	14	31.8	47	1	US-08-171-389-250	Sequence 250, App
	42	14	31.8	47	1	US-08-123-936-250	Sequence 250, App
	43	14	31.8	47	2	US-08-475-228A-250	Sequence 250, App
	44	14	31.8	47	3	US-08-482-080A-250	Sequence 250, App
	45	14	31.8	47	3	US-09-354-947-250	Sequence 250, App
	46	14	31.8	47	4	US-09-422-978-990	Sequence 990, App
	47	14	31.8	47	5	PCT-US93-12388-250	Sequence 250, App
	48	13.8	31.4	25	4	US-09-396-196G-50594	Sequence 50594, A
	49	13.8	31.4	26	1	US-08-218-933-2	Sequence 2, Appli
	50	13.8	31.4	26	5	PCT-US95-03918-2	Sequence 2, Appli
	51	13.8	31.4	29	3	US-08-194-560-7	Sequence 7, Appli
	52	13.8	31.4	34	3	US-08-510-133A-7	Sequence 7, Appli
	53	13.8	31.4	34	3	US-08-585-895-7	Sequence 7, Appli
	54	13.8	31.4	34	3	US-08-601-132-7	Sequence 7, Appli
	55	13.8	31.4	34	4	US-08-671-573B-7	Sequence 7, Appli
	56	13.8	31.4	34	4	US-09-631-092B-7	Sequence 7, Appli
c	57	13.8	31.4	38	3	US-08-718-904-124	Sequence 124, App
c	58	13.8	31.4	38	4	US-09-449-249-124	Sequence 124, App
c	59	13.8	31.4	47	4	US-09-516-667-73	Sequence 73, Appl
c	60	13.8	31.4	48	4	US-09-516-667-35	Sequence 35, Appl
c	61	13.8	31.4	48	4	US-09-516-667-74	Sequence 74, Appl
c	62	13.8	31.4	48	4	US-09-516-667-75	Sequence 75, Appl
	63	13.6	30.9	25	4	US-09-396-196G-84468	Sequence 84468, A
	64	13.6	30.9	25	4	US-09-396-196G-84469	Sequence 84469, A
c	65	13.6	30.9	38	4	US-09-060-299-249	Sequence 249, App
c	66	13.6	30.9	38	4	US-09-402-923A-249	Sequence 249, App
c	67	13.6	30.9	40	2	US-08-628-422-25	Sequence 25, Appl
	68	13.6	30.9	40	4	US-09-763-590-14	Sequence 14, Appl
	69	13.6	30.9	42	4	US-09-408-020-88	Sequence 88, Appl
	70	13.6	30.9	42	6	5177307-5	Patent No. 5177307
	71	13.6	30.9	42	6	5177307-5	Patent No. 5177307
c	72	13.4	30.5	33	1	US-08-449-311A-5	Sequence 5, Appli
c	73	13.4	30.5	33	5	PCT-US95-17106A-5	Sequence 5, Appli
c	74	13.4	30.5	36	1	US-08-121-202-21	Sequence 21, Appl
c	75	13.4	30.5	36	1	US-08-121-202-22	Sequence 22, Appl
c	76	13.4	30.5	36	3	US-08-822-516-8	Sequence 8, Appli
c	77	13.4	30.5	36	3	US-09-131-684-8	Sequence 8, Appli
	78	13.4	30.5	39	6	5519127-31	Patent No. 5519127
	79	13.4	30.5	39	6	5519127-31	Patent No. 5519127
	80	13.4	30.5	40	2	US-08-124-981A-20	Sequence 20, Appl
	81	13.4	30.5	40	3	US-09-037-190-18	Sequence 18, Appl
	82	13.4	30.5	40	3	US-09-037-192-18	Sequence 18, Appl
	83	13.4	30.5	40	3	US-09-037-143-18	Sequence 18, Appl
	84	13.4	30.5	40	3	US-09-049-691-18	Sequence 18, Appl
	85	13.4	30.5	40	3	US-08-260-174-18	Sequence 18, Appl
	86	13.4	30.5	40	3	US-09-338-128A-18	Sequence 18, Appl
	87	13.4	30.5	40	3	US-09-232-346-18	Sequence 18, Appl
	88	13.4	30.5	40	3	US-09-037-192-18	Sequence 18, Appl
c	89	13.4	30.5	45	3	US-08-358-627F-3	Sequence 3, Appli
c	90	13.4	30.5	45	3	US-08-465-712C-3	Sequence 3, Appli
c	91	13.4	30.5	45	3	US-09-552-733-3	Sequence 3, Appli
c	92	13.4	30.5	45	4	US-09-349-925-3	Sequence 3, Appli
	93	13.4	30.5	49	3	US-08-782-480-57	Sequence 57, Appl
	94	13.4	30.5	49	3	US-08-954-211-57	Sequence 57, Appl
	95	13.4	30.5	49	3	US-09-005-167A-57	Sequence 57, Appl
	96	13.4	30.5	49	3	US-09-176-741B-57	Sequence 57, Appl
	97	13.4	30.5	50	1	US-07-750-080A-21	Sequence 21, Appl
	98	13.4	30.5	50	1	US-07-750-080A-42	Sequence 42, Appl
	99	13.4	30.5	50	3	US-08-651-472-21	Sequence 21, Appl
	100	13.4	30.5	50	3	US-08-651-472-42	Sequence 42, Appl

ALIGNMENTS

```
RESULT 1
US-09-396-196G-29372/c
; Sequence 29372, Application US/09396196G
; Patent No. 6821724
; GENERAL INFORMATION:
;  APPLICANT: Michael Mittmann
;  APPLICANT:  David Mack
;  APPLICANT:  David Lockhart
;  APPLICANT:  Affymetrix, Inc.
;  TITLE OF INVENTION: Methods of Genetic Analysis
;  FILE REFERENCE: 3101.1
;  CURRENT APPLICATION NUMBER: US/09/396,196G
;  CURRENT FILING DATE:  1999-09-15
;  PRIOR APPLICATION NUMBER: 60/100,678
;  PRIOR FILING DATE: 1998-09-17
;  NUMBER OF SEQ ID NOS: 127806
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 29372
;   LENGTH: 25
;   TYPE: DNA
;   ORGANISM: Mus musculus
US-09-396-196G-29372

  Query Match             35.9%;  Score 15.8;  DB 4;  Length 25;
  Best Local Similarity   89.5%;  Pred. No. 1.5e+03;
  Matches   17;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy         26 CGGTCGCGGTTATTAAGAA 44
              | |||||||| |||||||
Db         19 CAGTCGCGGTAATTAAGAA 1


RESULT 2
5519127-36
;Patent No. 5519127
;    APPLICANT: SHAH, JYOTSNA;BUHARIN, AMELIA;LANE, DAVID J.
;    TITLE OF INVENTION: NUCLEIC ACID PROBES FOR THE DETECTION OF
;PNEUMOCYSTIS CARINII
;    NUMBER OF SEQUENCES: 57
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/826,657
;      FILING DATE: 21-JAN-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER: 392,679
;      FILING DATE: 11-AUG-1989
;SEQ ID NO:36:
;      LENGTH: 33
5519127-36

  Query Match             34.5%;  Score 15.2;  DB 6;  Length 33;
  Best Local Similarity   45.5%;  Pred. No. 3e+03;
  Matches   10;  Conservative    8;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          8 CCGTTCCTTCTTAATAACCGGT 29
              :| ::||::|:  |: |||||:
Db          3 YCCUUCCUUCUGGAUUACCGGU 24


RESULT 3
5519127-36
;Patent No. 5519127
;    APPLICANT: SHAH, JYOTSNA;BUHARIN, AMELIA;LANE, DAVID J.
;    TITLE OF INVENTION: NUCLEIC ACID PROBES FOR THE DETECTION OF
;PNEUMOCYSTIS CARINII
;    NUMBER OF SEQUENCES: 57
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/826,657
;      FILING DATE: 21-JAN-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER: 392,679
;      FILING DATE: 11-AUG-1989
;SEQ ID NO:36:
;      LENGTH: 33
5519127-36

  Query Match             34.5%;  Score 15.2;  DB 6;  Length 33;
  Best Local Similarity   45.5%;  Pred. No. 3e+03;
  Matches   10;  Conservative    8;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          8 CCGTTCCTTCTTAATAACCGGT 29
              :| ::||::|:  |: |||||:
Db          3 YCCUUCCUUCUGGAUUACCGGU 24


RESULT 4
US-09-396-196G-29373/c
; Sequence 29373, Application US/09396196G
; Patent No. 6821724
; GENERAL INFORMATION:
;  APPLICANT: Michael Mittmann
;  APPLICANT:  David Mack
;  APPLICANT:  David Lockhart
;  APPLICANT:  Affymetrix, Inc.
;  TITLE OF INVENTION: Methods of Genetic Analysis
;  FILE REFERENCE: 3101.1
;  CURRENT APPLICATION NUMBER: US/09/396,196G
;  CURRENT FILING DATE:  1999-09-15
;  PRIOR APPLICATION NUMBER: 60/100,678
;  PRIOR FILING DATE: 1998-09-17
;  NUMBER OF SEQ ID NOS: 127806
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 29373
;   LENGTH: 25
;   TYPE: DNA
;   ORGANISM: Mus musculus
US-09-396-196G-29373

  Query Match             33.6%;  Score 14.8;  DB 4;  Length 25;
  Best Local Similarity   88.9%;  Pred. No. 4.1e+03;
  Matches   16;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy         26 CGGTCGCGGTTATTAAGA 43
              | |||||||| |||||||
Db         18 CAGTCGCGGTAATTAAGA 1


RESULT 5
US-09-665-189A-40
; Sequence 40, Application US/09665189A
; Patent No. 6645765
; GENERAL INFORMATION:
;  APPLICANT: Anderson, Heather
;  APPLICANT:  Chay, Catherine
;  APPLICANT:  Chen, Guilan
;  APPLICANT:  Conner, Timothy
;  TITLE OF INVENTION: Plant Regulatory Sequences for Control of Gene Expression
;  FILE REFERENCE: 38-211(15674)B
;  CURRENT APPLICATION NUMBER: US/09/665,189A
;  CURRENT FILING DATE:  2000-09-15
;  PRIOR APPLICATION NUMBER: 09/665,189
;  PRIOR FILING DATE: 2000-09-15
;  NUMBER OF SEQ ID NOS: 75
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 40
;   LENGTH: 36
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(36)
;   OTHER INFORMATION: synthetic primer sequence
```

```
US-09-665-189A-40

  Query Match             33.6%;  Score 14.8;  DB 4;  Length 36;
  Best Local Similarity   73.1%;  Pred. No. 4.5e+03;
  Matches   19;  Conservative    0;  Mismatches    7;  Indels    0;  Gaps    0;

Qy          9 CGTTCCTTCTTAATAACCGGTCGCGG 34
              | || ||||| | | | |||| ||||
Db          6 CCTTTCTTCTCACTCAGCGGTTGCGG 31


RESULT 6
US-09-422-978-240/c
; Sequence 240, Application US/09422978
; Patent No. 6537751
; GENERAL INFORMATION:
;  APPLICANT: Cohen, Daniel
;  APPLICANT: Blumenfeld, Marta
;  APPLICANT: Chumakov, Ilya
;  TITLE OF INVENTION: Biallelic markers for use in constructing a high density...
;  FILE REFERENCE: GENSET.020CP1
;  CURRENT APPLICATION NUMBER: US/09/422,978
;  CURRENT FILING DATE: 1999-10-20
;  EARLIER APPLICATION NUMBER: US 09/298,850
;  EARLIER FILING DATE: 1999-04-21
;  EARLIER APPLICATION NUMBER: US 60/109,732
;  EARLIER FILING DATE: 1998-11-23
;  EARLIER APPLICATION NUMBER: US 60/082,614
;  EARLIER FILING DATE: 1998-04-21
;  NUMBER OF SEQ ID NOS: 11796
; SEQ ID NO 240
;   LENGTH: 47
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
;   FEATURE:
;   NAME/KEY: allele
;   LOCATION: 24
;   OTHER INFORMATION: 99-1368-299 : polymorphic base C or T
US-09-422-978-240

  Query Match             33.6%;  Score 14.8;  DB 4;  Length 47;
  Best Local Similarity   61.1%;  Pred. No. 4.8e+03;
  Matches   22;  Conservative    1;  Mismatches   13;  Indels    0;  Gaps    0;

Qy          9 CGTTCCTTCTTAATAACCGGTCGCGGTTATTAAGAA 44
              | ||   | ||||||   | ||:| ||| | || ||
Db         46 CATTTAATTTTAATACATGCTCRCTGTTTTGAAAAA 11


RESULT 7
US-09-422-978-1725/c
; Sequence 1725, Application US/09422978
; Patent No. 6537751
; GENERAL INFORMATION:
;  APPLICANT: Cohen, Daniel
;  APPLICANT: Blumenfeld, Marta
;  APPLICANT: Chumakov, Ilya
;  TITLE OF INVENTION: Biallelic markers for use in constructing a high density...
;  FILE REFERENCE: GENSET.020CP1
;  CURRENT APPLICATION NUMBER: US/09/422,978
;  CURRENT FILING DATE: 1999-10-20
;  EARLIER APPLICATION NUMBER: US 09/298,850
;  EARLIER FILING DATE: 1999-04-21
;  EARLIER APPLICATION NUMBER: US 60/109,732
;  EARLIER FILING DATE: 1998-11-23
;  EARLIER APPLICATION NUMBER: US 60/082,614
;  EARLIER FILING DATE: 1998-04-21
;  NUMBER OF SEQ ID NOS: 11796
; SEQ ID NO 1725
;   LENGTH: 47
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
;   FEATURE:
;   NAME/KEY: allele
;   LOCATION: 24
;   OTHER INFORMATION: 99-5951-438 : polymorphic base C or T
US-09-422-978-1725

  Query Match             33.6%;  Score 14.8;  DB 4;  Length 47;
  Best Local Similarity   59.5%;  Pred. No. 4.8e+03;
  Matches   25;  Conservative    0;  Mismatches   17;  Indels    0;  Gaps    0;

Qy          3 GGGTCCCGTTCCTTCTTAATAACCGGTCGCGGTTATTAAGAA 44
              || ||||  ||  |||| ||||     | |  ||| | | ||
Db         46 GGCTCCCCATCTCTCTTTATAARTTAGCACCATTAATTATAA 5


RESULT 8
US-08-171-389-585/c
; Sequence 585, Application US/08171389
; Patent No. 5578444
;  GENERAL INFORMATION:
;    APPLICANT:  Edwards, Cynthia A.
;    APPLICANT:  Cantor, Charles R.
;    APPLICANT:  Andrews, Beth M.
;    APPLICANT:  Turin, Lisa M.
;    APPLICANT:  Fry, Kirk E.
;    TITLE OF INVENTION:  Sequence-Directed DNA Binding
;    TITLE OF INVENTION:  Molecules, Compositions and Methods
;    NUMBER OF SEQUENCES:  641
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Genelabs Technologies, Inc.
;      STREET:  505 Penobscot Drive
;      CITY:  Redwood City
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94063
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/171,389
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/123,936
;      FILING DATE:  17-SEP-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/996,783
;      FILING DATE:  23-DEC-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/723,618
;      FILING DATE:  27-JUN-1991
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/081,070
;      FILING DATE:  22-JUN-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Fabian, Gary R.
;      REGISTRATION NUMBER:  33,875
;      REFERENCE/DOCKET NUMBER:  4600-0175/G19P3
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 324-0880
;      TELEFAX:  (415) 324-0960
;  INFORMATION FOR SEQ ID NO:  585:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  50 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    HYPOTHETICAL:  NO
;    ORIGINAL SOURCE:
;      INDIVIDUAL ISOLATE:  Human papilloma virus type-16 E6/E7
;      INDIVIDUAL ISOLATE:  (start site 97)
```

US-08-171-389-585

Query Match 33.6%; Score 14.8; DB 1; Length 50;
Best Local Similarity 73.1%; Pred. No. 4.9e+03;
Matches 19; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

```
Qy         11 TTCCTTCTTAATAACCGGTCGCGGTT 36
              | | ||  || ||||||||  |||||
Db         27 TGCTTTTATACTAACCGGTTTCGGTT 2
```

RESULT 9
US-08-123-936-585/c
; Sequence 585, Application US/08123936
; Patent No. 5726014
; GENERAL INFORMATION:
; APPLICANT: Edwards, Cynthia A.
; APPLICANT: Cantor, Charles R.
; APPLICANT: Andrews, Beth M.
; APPLICANT: Turin, Lisa M.
; TITLE OF INVENTION: Screening Assay for the Detection of
; TITLE OF INVENTION: DNA-Binding Molecules
; NUMBER OF SEQUENCES: 640
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genelabs Technologies, Inc.
; STREET: 505 Penobscot Drive
; CITY: Redwood City
; STATE: CA
; COUNTRY: USA
; ZIP: 94063
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/123,936
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/996,783
; FILING DATE: 23-DEC-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/723,618
; FILING DATE: 27-JUN-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Fabian, Gary R.
; REGISTRATION NUMBER: 33,875
; REFERENCE/DOCKET NUMBER: 4600-0075.32/G19P2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 324-0880
; TELEFAX: (415) 324-0960
; INFORMATION FOR SEQ ID NO: 585:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 50 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; INDIVIDUAL ISOLATE: Human papilloma virus type-16 E6/E7
; INDIVIDUAL ISOLATE: (start site 97)
US-08-123-936-585

Query Match 33.6%; Score 14.8; DB 1; Length 50;
Best Local Similarity 73.1%; Pred. No. 4.9e+03;
Matches 19; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

```
Qy         11 TTCCTTCTTAATAACCGGTCGCGGTT 36
              | | ||  || ||||||||  |||||
Db         27 TGCTTTTATACTAACCGGTTTCGGTT 2
```

RESULT 10
US-08-475-228A-585/c
; Sequence 585, Application US/08475228A
; Patent No. 5869241
; GENERAL INFORMATION:
; APPLICANT: Edwards, Cynthia A.
; APPLICANT: Cantor, Charles R.
; APPLICANT: Andrews, Beth M.
; APPLICANT: Turin, Lisa M.
; APPLICANT: Fry, Kirk E.
; TITLE OF INVENTION: Sequence-Directed DNA Binding
; TITLE OF INVENTION: Molecules, Compositions and Methods
; NUMBER OF SEQUENCES: 664
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genelabs Technologies, Inc.
; STREET: 505 Penobscot Drive
; CITY: Redwood City
; STATE: CA
; COUNTRY: USA
; ZIP: 94063
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/475,228A
; FILING DATE: 06-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/123,936
; FILING DATE: 17-SEP-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/996,783
; FILING DATE: 23-DEC-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/723,618
; FILING DATE: 27-JUN-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/081,070
; FILING DATE: 22-JUN-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Stratford, Carol A.
; REGISTRATION NUMBER: 34,444
; REFERENCE/DOCKET NUMBER: 4600-0175.21/G19P3D2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 324-0880
; TELEFAX: (415) 324-0960
; INFORMATION FOR SEQ ID NO: 585:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 50 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; INDIVIDUAL ISOLATE: Human papilloma virus type-16 E6/E7
; INDIVIDUAL ISOLATE: (start site 97)
US-08-475-228A-585

Query Match 33.6%; Score 14.8; DB 2; Length 50;
Best Local Similarity 73.1%; Pred. No. 4.9e+03;
Matches 19; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

```
Qy         11 TTCCTTCTTAATAACCGGTCGCGGTT 36
              | | ||  || ||||||||  |||||
Db         27 TGCTTTTATACTAACCGGTTTCGGTT 2
```

RESULT 11
US-08-482-080A-585/c
; Sequence 585, Application US/08482080A

```
; Patent No. 6010849
;  GENERAL INFORMATION:
;    APPLICANT:  Edwards, Cynthia A.
;    APPLICANT:  Cantor, Charles R.
;    APPLICANT:  Andrews, Beth M.
;    APPLICANT:  Turin, Lisa M.
;    APPLICANT:  Fry, Kirk E.
;    TITLE OF INVENTION:  Sequence-Directed DNA Binding
;    TITLE OF INVENTION:  Molecules, Compositions and Methods
;    NUMBER OF SEQUENCES:  664
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Genelabs Technologies, Inc.
;      STREET:  505 Penobscot Drive
;      CITY:  Redwood City
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94063
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/482,080A
;      FILING DATE:  07-JUN-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/171,389
;      FILING DATE:  20-DEC-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/123,936
;      FILING DATE:  17-SEP-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/996,783
;      FILING DATE:  23-DEC-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/723,618
;      FILING DATE:  27-JUN-1991
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/081,070
;      FILING DATE:  22-JUN-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Brady, John F.
;      REGISTRATION NUMBER:  39,118
;      REFERENCE/DOCKET NUMBER:  4600-0175.20/G19P3D1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (650) 324-0880
;      TELEFAX:  (650) 324-0960
;  INFORMATION FOR SEQ ID NO:  585:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  50 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    HYPOTHETICAL:  NO
;    ORIGINAL SOURCE:
;      INDIVIDUAL ISOLATE:  Human papilloma virus type-16 E6/E7
;      INDIVIDUAL ISOLATE:  (start site 97)
US-08-482-080A-585

  Query Match             33.6%;  Score 14.8;  DB 3;  Length 50;
  Best Local Similarity   73.1%;  Pred. No. 4.9e+03;
  Matches   19;  Conservative    0;  Mismatches    7;  Indels    0;  Gaps    0;

Qy         11 TTCCTTCTTAATAACCGGTCGCGGTT 36
              | | ||  || ||||||||  |||||
Db         27 TGCTTTTATACTAACCGGTTTCGGTT 2


RESULT 12
US-09-354-947-585/c
; Sequence 585, Application US/09354947
; Patent No. 6384208
;  GENERAL INFORMATION:
;    APPLICANT:  Edwards, Cynthia A.
;    APPLICANT:  Cantor, Charles R.
;    APPLICANT:  Andrews, Beth M.
;    APPLICANT:  Turin, Lisa M.
;    APPLICANT:  Fry, Kirk E.
;    TITLE OF INVENTION:  Sequence-Directed DNA Binding
;    TITLE OF INVENTION:  Molecules, Compositions and Methods
;    NUMBER OF SEQUENCES:  664
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Genelabs Technologies, Inc.
;      STREET:  505 Penobscot Drive
;      CITY:  Redwood City
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94063
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/354,947
;      FILING DATE:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/482,080
;      FILING DATE:  07-JUN-1995
;      APPLICATION NUMBER:  US 08/171,389
;      FILING DATE:  20-DEC-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/123,936
;      FILING DATE:  17-SEP-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/996,783
;      FILING DATE:  23-DEC-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/723,618
;      FILING DATE:  27-JUN-1991
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/081,070
;      FILING DATE:  22-JUN-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Brady, John F.
;      REGISTRATION NUMBER:  39,118
;      REFERENCE/DOCKET NUMBER:  4600-0175.20/G19P3D1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (650) 324-0880
;      TELEFAX:  (650) 324-0960
;  INFORMATION FOR SEQ ID NO:  585:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  50 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    HYPOTHETICAL:  NO
;    ORIGINAL SOURCE:
;      INDIVIDUAL ISOLATE:  Human papilloma virus type-16 E6/E7
;      INDIVIDUAL ISOLATE:  (start site 97)
US-09-354-947-585

  Query Match             33.6%;  Score 14.8;  DB 3;  Length 50;
  Best Local Similarity   73.1%;  Pred. No. 4.9e+03;
  Matches   19;  Conservative    0;  Mismatches    7;  Indels    0;  Gaps    0;

Qy         11 TTCCTTCTTAATAACCGGTCGCGGTT 36
              | | ||  || ||||||||  |||||
Db         27 TGCTTTTATACTAACCGGTTTCGGTT 2


RESULT 13
PCT-US93-12388-585/c
; Sequence 585, Application PC/TUS9312388
```

```
;  GENERAL INFORMATION:
;    APPLICANT:
;    TITLE OF INVENTION:  Sequence-Directed DNA Binding
;    TITLE OF INVENTION:  Molecules, Compositions and Methods
;    NUMBER OF SEQUENCES:  641
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Genelabs Technologies, Inc.
;      STREET:  505 Penobscot Drive
;      CITY:  Redwood City
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94063
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US93/12388
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/123,936
;      FILING DATE:  17-SEP-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/996,783
;      FILING DATE:  23-DEC-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Fabian, Gary R.
;      REGISTRATION NUMBER:  33,875
;      REFERENCE/DOCKET NUMBER:  4600-0175.41/G19PCT2
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 324-0880
;      TELEFAX:  (415) 324-0960
;  INFORMATION FOR SEQ ID NO:  585:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  50 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    HYPOTHETICAL:  NO
;    ORIGINAL SOURCE:
;      INDIVIDUAL ISOLATE:  Human papilloma virus type-16 E6/E7
;      INDIVIDUAL ISOLATE:  (start site 97)
PCT-US93-12388-585

  Query Match             33.6%;  Score 14.8;  DB 5;  Length 50;
  Best Local Similarity   73.1%;  Pred. No. 4.9e+03;
  Matches   19;  Conservative    0;  Mismatches    7;  Indels    0;  Gaps    0;

Qy         11 TTCCTTCTTAATAACCGGTCGCGGTT 36
              | | ||  || ||||||||  |||||
Db         27 TGCTTTTATACTAACCGGTTTCGGTT 2


RESULT 14
5519127-17
;Patent No. 5519127
;    APPLICANT: SHAH, JYOTSNA;BUHARIN, AMELIA;LANE, DAVID J.
;    TITLE OF INVENTION: NUCLEIC ACID PROBES FOR THE DETECTION OF
;PNEUMOCYSTIS CARINII
;    NUMBER OF SEQUENCES: 57
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/826,657
;      FILING DATE: 21-JAN-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER: 392,679
;      FILING DATE: 11-AUG-1989
;SEQ ID NO:17:
;      LENGTH: 33
5519127-17

  Query Match             33.2%;  Score 14.6;  DB 6;  Length 33;
  Best Local Similarity   81.0%;  Pred. No. 5.4e+03;
  Matches   17;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          9 CGTTCCTTCTTAATAACCGGT 29
              | ||||||||  || ||||||
Db          4 CCTTCCTTCTGGATTACCGGT 24


RESULT 15
5519127-17
;Patent No. 5519127
;    APPLICANT: SHAH, JYOTSNA;BUHARIN, AMELIA;LANE, DAVID J.
;    TITLE OF INVENTION: NUCLEIC ACID PROBES FOR THE DETECTION OF
;PNEUMOCYSTIS CARINII
;    NUMBER OF SEQUENCES: 57
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/826,657
;      FILING DATE: 21-JAN-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER: 392,679
;      FILING DATE: 11-AUG-1989
;SEQ ID NO:17:
;      LENGTH: 33
5519127-17

  Query Match             33.2%;  Score 14.6;  DB 6;  Length 33;
  Best Local Similarity   81.0%;  Pred. No. 5.4e+03;
  Matches   17;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          9 CGTTCCTTCTTAATAACCGGT 29
              | ||||||||  || ||||||
Db          4 CCTTCCTTCTGGATTACCGGT 24


RESULT 16
5519127-4/c
;Patent No. 5519127
;    APPLICANT: SHAH, JYOTSNA;BUHARIN, AMELIA;LANE, DAVID J.
;    TITLE OF INVENTION: NUCLEIC ACID PROBES FOR THE DETECTION OF
;PNEUMOCYSTIS CARINII
;    NUMBER OF SEQUENCES: 57
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/826,657
;      FILING DATE: 21-JAN-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER: 392,679
;      FILING DATE: 11-AUG-1989
;SEQ ID NO:4:
;      LENGTH: 34
5519127-4

  Query Match             33.2%;  Score 14.6;  DB 6;  Length 34;
  Best Local Similarity   81.0%;  Pred. No. 5.4e+03;
  Matches   17;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          9 CGTTCCTTCTTAATAACCGGT 29
              | ||||||||  || ||||||
Db         31 CCTTCCTTCTGGATTACCGGT 11


RESULT 17
5519127-4/c
;Patent No. 5519127
;    APPLICANT: SHAH, JYOTSNA;BUHARIN, AMELIA;LANE, DAVID J.
;    TITLE OF INVENTION: NUCLEIC ACID PROBES FOR THE DETECTION OF
;PNEUMOCYSTIS CARINII
;    NUMBER OF SEQUENCES: 57
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/826,657
;      FILING DATE: 21-JAN-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER: 392,679
```

```
;      FILING DATE: 11-AUG-1989
;SEQ ID NO:4:
;      LENGTH: 34
5519127-4

  Query Match             33.2%;  Score 14.6;  DB 6;  Length 34;
  Best Local Similarity   81.0%;  Pred. No. 5.4e+03;
  Matches   17;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          9 CGTTCCTTCTTAATAACCGGT 29
              | ||||||||  || ||||||
Db         31 CCTTCCTTCTGGATTACCGGT 11


RESULT 18
5519127-33
;Patent No. 5519127
;    APPLICANT: SHAH, JYOTSNA;BUHARIN, AMELIA;LANE, DAVID J.
;    TITLE OF INVENTION: NUCLEIC ACID PROBES FOR THE DETECTION OF
;PNEUMOCYSTIS CARINII
;    NUMBER OF SEQUENCES: 57
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/826,657
;      FILING DATE: 21-JAN-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER: 392,679
;      FILING DATE: 11-AUG-1989
;SEQ ID NO:33:
;      LENGTH: 39
5519127-33

  Query Match             33.2%;  Score 14.6;  DB 6;  Length 39;
  Best Local Similarity   47.6%;  Pred. No. 5.6e+03;
  Matches   10;  Conservative    7;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          9 CGTTCCTTCTTAATAACCGGT 29
              | ::||::|:  |: |||||:
Db          5 CCUUCCUUCUGGAUUACCGGU 25


RESULT 19
5519127-33
;Patent No. 5519127
;    APPLICANT: SHAH, JYOTSNA;BUHARIN, AMELIA;LANE, DAVID J.
;    TITLE OF INVENTION: NUCLEIC ACID PROBES FOR THE DETECTION OF
;PNEUMOCYSTIS CARINII
;    NUMBER OF SEQUENCES: 57
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/826,657
;      FILING DATE: 21-JAN-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER: 392,679
;      FILING DATE: 11-AUG-1989
;SEQ ID NO:33:
;      LENGTH: 39
5519127-33

  Query Match             33.2%;  Score 14.6;  DB 6;  Length 39;
  Best Local Similarity   47.6%;  Pred. No. 5.6e+03;
  Matches   10;  Conservative    7;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          9 CGTTCCTTCTTAATAACCGGT 29
              | ::||::|:  |: |||||:
Db          5 CCUUCCUUCUGGAUUACCGGU 25


RESULT 20
US-07-826-928A-30
; Sequence 30, Application US/07826928A
; Patent No. 5439829
;  GENERAL INFORMATION:
;    APPLICANT:  Anderson, Leslie D.
;    APPLICANT:  Cook, James A.
;    APPLICANT:  David, Gary S.
;    APPLICANT:  Hochschwender, Susan M.
;    APPLICANT:  Kasher, Mary S.
;    APPLICANT:  Smith, Michele C.
;    APPLICANT:  Stemmer, Willem P.
;    TITLE OF INVENTION:  METHOD OF IMMOBILIZING AND CROSS LINKING
;    TITLE OF INVENTION:  PROTEINS AND OTHER MOLECULES AND USES THEREOF
;    NUMBER OF SEQUENCES:  38
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Eli Lilly and Company
;      STREET:  Lilly Corporate Center
;      CITY:  Indianapolis
;      STATE:  IN
;      COUNTRY:  USA
;      ZIP:  46285
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/826,928A
;      FILING DATE:  19920124
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Murphy, Richard B.
;      REGISTRATION NUMBER:  35,296
;      REFERENCE/DOCKET NUMBER:  X8180A
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  317-276-3589
;      TELEFAX:  317-276-1294
;  INFORMATION FOR SEQ ID NO:  30:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  46 base pairs
;      TYPE:  NUCLEIC ACID
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-07-826-928A-30

  Query Match             33.2%;  Score 14.6;  DB 1;  Length 46;
  Best Local Similarity   69.0%;  Pred. No. 5.9e+03;
  Matches   20;  Conservative    0;  Mismatches    9;  Indels    0;  Gaps    0;

Qy         16 TCTTAATAACCGGTCGCGGTTATTAAGAA 44
              | |||||||    ||||  |||| ||| |
Db          9 TATTAATAATGTATCGCTATTATAAAGGA 37


RESULT 21
US-09-422-978-3415
; Sequence 3415, Application US/09422978
; Patent No. 6537751
; GENERAL INFORMATION:
;  APPLICANT: Cohen, Daniel
;  APPLICANT: Blumenfeld, Marta
;  APPLICANT: Chumakov, Ilya
;  TITLE OF INVENTION: Biallelic markers for use in constructing a high density...
;  FILE REFERENCE: GENSET.020CP1
;  CURRENT APPLICATION NUMBER: US/09/422,978
;  CURRENT FILING DATE: 1999-10-20
;  EARLIER APPLICATION NUMBER: US 09/298,850
;  EARLIER FILING DATE: 1999-04-21
;  EARLIER APPLICATION NUMBER: US 60/109,732
;  EARLIER FILING DATE: 1998-11-23
;  EARLIER APPLICATION NUMBER: US 60/082,614
;  EARLIER FILING DATE: 1998-04-21
;  NUMBER OF SEQ ID NOS: 11796
; SEQ ID NO 3415
;   LENGTH: 47
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
;   FEATURE:
```

```
;   NAME/KEY: allele
;   LOCATION: 24
;   OTHER INFORMATION: 99-3812-243 : polymorphic base T or G
US-09-422-978-3415

  Query Match             33.2%;  Score 14.6;  DB 4;  Length 47;
  Best Local Similarity   64.5%;  Pred. No. 5.9e+03;
  Matches   20;  Conservative    1;  Mismatches   10;  Indels    0;  Gaps    0;

Qy         10 GTTCCTTCTTAATAACCGGTCGCGGTTATTA 40
              |||| | |:||| ||||  || | | |  ||
Db         16 GTTCATCCKTAAAAACCATTCTCAGCTCCTA 46


RESULT 22
US-08-889-502-34
; Sequence 34, Application US/08889502
; Patent No. 6066726
;  GENERAL INFORMATION:
;    APPLICANT:  Farb, David H
;    APPLICANT:  Russek, Shelley J
;    TITLE OF INVENTION:  GENE THERAPY VECTOR WITH TISSUE
;    TITLE OF INVENTION:  SPECIFICITY
;    NUMBER OF SEQUENCES:  37
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Kevin M. Farrell
;      STREET:  P.O. Box 999
;      CITY:  York Harbor
;      STATE:  ME
;      COUNTRY:  USA
;      ZIP:  03911
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/889,502
;      FILING DATE:  08-JUL-1997
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Farrell, Kevin M
;      REGISTRATION NUMBER:  35,505
;      REFERENCE/DOCKET NUMBER:  0146-2008
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (207) 363-0558
;      TELEFAX:  (207) 363-0528
;  INFORMATION FOR SEQ ID NO:  34:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  49 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-08-889-502-34

  Query Match             33.2%;  Score 14.6;  DB 3;  Length 49;
  Best Local Similarity   62.2%;  Pred. No. 6e+03;
  Matches   23;  Conservative    0;  Mismatches   14;  Indels    0;  Gaps    0;

Qy          3 GGGTCCCGTTCCTTCTTAATAACCGGTCGCGGTTATT 39
              | ||| | |||||||| | |   | |  ||  ||| |
Db         10 GTGTCGCTTTCCTTCTCACTTGTCAGGGGCTCTTAGT 46


RESULT 23
US-09-422-978-2970
; Sequence 2970, Application US/09422978
; Patent No. 6537751
; GENERAL INFORMATION:
;  APPLICANT: Cohen, Daniel
;  APPLICANT: Blumenfeld, Marta
;  APPLICANT: Chumakov, Ilya
;  TITLE OF INVENTION: Biallelic markers for use in constructing a high density...
;  FILE REFERENCE: GENSET.020CP1
;  CURRENT APPLICATION NUMBER: US/09/422,978
;  CURRENT FILING DATE: 1999-10-20
;  EARLIER APPLICATION NUMBER: US 09/298,850
;  EARLIER FILING DATE: 1999-04-21
;  EARLIER APPLICATION NUMBER: US 60/109,732
;  EARLIER FILING DATE: 1998-11-23
;  EARLIER APPLICATION NUMBER: US 60/082,614
;  EARLIER FILING DATE: 1998-04-21
;  NUMBER OF SEQ ID NOS: 11796
; SEQ ID NO 2970
;   LENGTH: 47
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
;   FEATURE:
;   NAME/KEY: allele
;   LOCATION: 24
;   OTHER INFORMATION: 99-21307-370 : polymorphic base A or G
US-09-422-978-2970

  Query Match             32.7%;  Score 14.4;  DB 4;  Length 47;
  Best Local Similarity   69.2%;  Pred. No. 7.2e+03;
  Matches   18;  Conservative    1;  Mismatches    7;  Indels    0;  Gaps    0;

Qy          1 GCGGGTCCCGTTCCTTCTTAATAACC 26
              ||:  ||| ||| | | |||||||  |
Db         22 GCRCCTCCTGTTGCCTTTTAATAAGC 47


RESULT 24
US-09-198-452A-6719/c
; Sequence 6719, Application US/09198452A
; Patent No. 6559294
; GENERAL INFORMATION:
;  APPLICANT: Griffais, R.
;  TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragments
;  TITLE OF INVENTION:  thereof and uses thereof, in particular for the diagnosis, preve
;  TITLE OF INVENTION:  and treatment of infection
;  FILE REFERENCE: 9710-003-999
;  CURRENT APPLICATION NUMBER: US/09/198,452A
;  CURRENT FILING DATE:  1998-11-24
;  NUMBER OF SEQ ID NOS: 6849
; SEQ ID NO 6719
;   LENGTH: 20
;   TYPE: DNA
;   ORGANISM: Chlamydia pneumoniae
US-09-198-452A-6719

  Query Match             32.3%;  Score 14.2;  DB 4;  Length 20;
  Best Local Similarity   84.2%;  Pred. No. 7.1e+03;
  Matches   16;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy         10 GTTCCTTCTTAATAACCGG 28
              |||||||||  ||||| ||
Db         19 GTTCCTTCTGGATAACAGG 1


RESULT 25
US-09-922-146-31/c
; Sequence 31, Application US/09922146
; Patent No. 6566133
; GENERAL INFORMATION:
;  APPLICANT: Lex M. Cowsert
;  APPLICANT:  Brett P. Monia
;  TITLE OF INVENTION: ANTISENSE MODULATION OF DUAL SPECIFIC PHOSPHATASE 9 EXPRESSION
;  FILE REFERENCE: RTS-0252
;  CURRENT APPLICATION NUMBER: US/09/922,146
;  CURRENT FILING DATE:  2001-08-01
;  NUMBER OF SEQ ID NOS: 48
; SEQ ID NO 31
;   LENGTH: 20
;   TYPE: DNA
```

```
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Antisense Oligonucleotide
US-09-922-146-31

  Query Match            32.3%;  Score 14.2;  DB 4;  Length 20;
  Best Local Similarity  84.2%;  Pred. No. 7.1e+03;
  Matches   16;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy         21 ATAACCGGTCGCGGTTATT 39
              | |||| ||||||||| ||
Db         19 AGAACCTGTCGCGGTTCTT 1


RESULT 26
US-09-657-472-1129/c
; Sequence 1129, Application US/09657472
; Patent No. 6727063
; GENERAL INFORMATION:
;  APPLICANT: Lander, Eric S.
;  APPLICANT:  Cargill, Michele
;  APPLICANT:  Ireland, James S.
;  APPLICANT:  Bolk, Stacey
;  APPLICANT:  Daley, George Q.
;  APPLICANT:  McCarthy, Jeanette J.
;  TITLE OF INVENTION: SINGLE NUCLEOTIDE POLYMORPHISMS IN GENES
;  FILE REFERENCE: 2825.1027-001
;  CURRENT APPLICATION NUMBER: US/09/657,472
;  CURRENT FILING DATE:  2000-09-07
;  PRIOR APPLICATION NUMBER: US 60/153,357
;  PRIOR FILING DATE: 1999-09-10
;  PRIOR APPLICATION NUMBER: US 60/220,947
;  PRIOR FILING DATE: 2000-07-26
;  PRIOR APPLICATION NUMBER: US 60/225,724
;  PRIOR FILING DATE: 2000-08-16
;  NUMBER OF SEQ ID NOS: 2551
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1129
;   LENGTH: 21
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-657-472-1129

  Query Match            32.3%;  Score 14.2;  DB 4;  Length 21;
  Best Local Similarity  76.2%;  Pred. No. 7.2e+03;
  Matches   16;  Conservative    1;  Mismatches    4;  Indels    0;  Gaps    0;

Qy         13 CCTTCTTAATAACCGGTCGCG 33
              | ||||||||:| | |||| |
Db         21 CTTTCTTAATRAGCTGTCGGG 1


RESULT 27
US-07-956-697B-3
; Sequence 3, Application US/07956697B
; Patent No. 5374543
;  GENERAL INFORMATION:
;    APPLICANT:  Murdock, Douglas Craig
;    TITLE OF INVENTION:  Enhanced Indole Biosynthesis
;    NUMBER OF SEQUENCES:  5
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Amgen Inc.
;      STREET:  Amgen Center
;      STREET:  1840 Dehavilland Drive
;      CITY:  Thousand Oaks
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  91320-1789
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  Microsoft Word Version 5.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/956,697B
;      FILING DATE:  02-OCT-1992
;      CLASSIFICATION:  435
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  36 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  unknown
US-07-956-697B-3

  Query Match            32.3%;  Score 14.2;  DB 1;  Length 36;
  Best Local Similarity  84.2%;  Pred. No. 8.2e+03;
  Matches   16;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy         25 CCGGTCGCGGTTATTAAGA 43
              |||||||||| || ||||
Db         11 CCGGTCGCGGCGATAAAGA 29


RESULT 28
US-08-263-098-3
; Sequence 3, Application US/08263098
; Patent No. 5494816
;  GENERAL INFORMATION:
;    APPLICANT:  Murdock, Douglas Craig
;    TITLE OF INVENTION:  Enhanced Indole Biosynthesis
;    NUMBER OF SEQUENCES:  5
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Amgen Inc.
;      STREET:  Amgen Center
;      STREET:  1840 Dehavilland Drive
;      CITY:  Thousand Oaks
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  91320-1789
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  Microsoft Word Version 5.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/263,098
;      FILING DATE:  21-JUN-1994
;      CLASSIFICATION:  435
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  36 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  unknown
US-08-263-098-3

  Query Match            32.3%;  Score 14.2;  DB 1;  Length 36;
  Best Local Similarity  84.2%;  Pred. No. 8.2e+03;
  Matches   16;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy         25 CCGGTCGCGGTTATTAAGA 43
              |||||||||| || ||||
Db         11 CCGGTCGCGGCGATAAAGA 29


RESULT 29
US-08-857-946-83/c
; Sequence 83, Application US/08857946
; Patent No. 5994075
;  GENERAL INFORMATION:
;    APPLICANT:  Goodfellow, P.N.
;    TITLE OF INVENTION:  METHODS FOR IDENTIFYING A MUTATION IN A
;    TITLE OF INVENTION:  GENE OF INTEREST
;    NUMBER OF SEQUENCES:  162
;    CORRESPONDENCE ADDRESS:
```

```
;      ADDRESSEE:  Banner & Witcoff, Inc.
;      STREET:  75 State Street
;      CITY:  Boston
;      STATE:  Massachusetts
;      COUNTRY:  USA
;      ZIP:  02109-1807
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  WordPerfect 6.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/857,946
;      FILING DATE:  16-MAY-1997
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US/60/017,824
;      FILING DATE:  17-MAY-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Kathleen M. Williams
;      REGISTRATION NUMBER:  34,380
;      REFERENCE/DOCKET NUMBER:  3529/05573
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617-345-9100
;      TELEFAX:  617-345-9111
;  INFORMATION FOR SEQ ID NO:  83:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  38 bases
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  other nucleic acid
;    FEATURE:
;      NAME/KEY:  primer bMm160f
US-08-857-946-83

  Query Match             32.3%;  Score 14.2;  DB 2;  Length 38;
  Best Local Similarity   62.9%;  Pred. No. 8.3e+03;
  Matches   22;  Conservative    0;  Mismatches   13;  Indels    0;  Gaps    0;

Qy          4 GGTCCCGTTCCTTCTTAATAACCGGTCGCGGTTAT 38
              ||||||  |   ||||   ||| || ||  ||| |
Db         38 GGTCCCAGTATGTCTTCTAAACTGGCCGTCGTTTT 4


RESULT 30
US-08-970-740-83/c
; Sequence 83, Application US/08970740
; Patent No. 6015670
;  GENERAL INFORMATION:
;    APPLICANT:  Goodfellow, P.N.
;    TITLE OF INVENTION:  METHODS FOR IDENTIFYING A MUTATION IN A
;    TITLE OF INVENTION:  GENE OF INTEREST
;    NUMBER OF SEQUENCES:  162
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Banner & Witcoff, Inc.
;      STREET:  28 State Street, 28th Floor
;      CITY:  Boston
;      STATE:  Massachusetts
;      COUNTRY:  USA
;      ZIP:  02109
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  WordPerfect 6.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/970,740
;      FILING DATE:  14-NOV-1997
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/857,946
;      FILING DATE:  16-MAY-1997
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  60/017,824
;      FILING DATE:  17-MAY-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Kathleen M. Williams
;      REGISTRATION NUMBER:  34,380
;      REFERENCE/DOCKET NUMBER:  3529/59829
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617-227-7111
;      TELEFAX:  617-227-4399
;  INFORMATION FOR SEQ ID NO:  83:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  38 bases
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  other nucleic acid
;    FEATURE:
;      NAME/KEY:  primer bMm160f
US-08-970-740-83

  Query Match             32.3%;  Score 14.2;  DB 3;  Length 38;
  Best Local Similarity   62.9%;  Pred. No. 8.3e+03;
  Matches   22;  Conservative    0;  Mismatches   13;  Indels    0;  Gaps    0;

Qy          4 GGTCCCGTTCCTTCTTAATAACCGGTCGCGGTTAT 38
              ||||||  |   ||||   ||| || ||  ||| |
Db         38 GGTCCCAGTATGTCTTCTAAACTGGCCGTCGTTTT 4


RESULT 31
US-09-538-709-983/c
; Sequence 983, Application US/09538709
; Patent No. 6468749
; GENERAL INFORMATION:
;  APPLICANT: Ulanovsky, et al
;  TITLE OF INVENTION: SEQUENCE-DEPENDENT GENE SORTING TECHNIQUES
;  FILE REFERENCE: 540579-2006
;  CURRENT APPLICATION NUMBER: US/09/538,709
;  CURRENT FILING DATE:  2001-06-08
;  NUMBER OF SEQ ID NOS: 1311
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 983
;   LENGTH: 49
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Adaptor
US-09-538-709-983

  Query Match             32.3%;  Score 14.2;  DB 3;  Length 49;
  Best Local Similarity   70.4%;  Pred. No. 8.9e+03;
  Matches   19;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Qy          9 CGTTCCTTCTTAATAACCGGTCGCGGT 35
              || || |||  |||   ||| ||||||
Db         44 CGCTCGTTCCGAATCGGCGGCCGCGGT 18


RESULT 32
US-08-379-452-17
; Sequence 17, Application US/08379452
; Patent No. 6040174
;  GENERAL INFORMATION:
;    APPLICANT:  IMLER, Jean-Luc
;    APPLICANT:  MEHTALI, Majid
;    APPLICANT:  PAVIRANI, Andrea
;    TITLE OF INVENTION:  DEFECTIVE ADENOVIRUSES AND CORRESPONDING
;    TITLE OF INVENTION:  COMPLEMENTATION LINES
;    NUMBER OF SEQUENCES:  43
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  BURNS, DOANE, SWECKER & MATHIS, L.L.P.
;      STREET:  1737 King Street, Suite 500
;      CITY:  Alexandria
```

```
;       STATE:  Virginia
;       COUNTRY:  United States
;       ZIP:  22314-2756
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.30
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/379,452
;       FILING DATE:  26-JAN-1995
;       CLASSIFICATION:  435
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  WO PCT/FR94/00624
;       FILING DATE:  27-MAY-1994
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  FR 93 06482
;       FILING DATE:  28-MAY-1993
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Dadio, Susan M.
;       REGISTRATION NUMBER:  40,373
;       REFERENCE/DOCKET NUMBER:  029395-002
;   INFORMATION FOR SEQ ID NO:  17:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  24 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  DNA (genomic)
;     HYPOTHETICAL:  NO
;     ANTI-SENSE:  NO
;     ORIGINAL SOURCE:
;       ORGANISM:  Synthetic oligonucleotide (OTG5482)
US-08-379-452-17

  Query Match             31.8%;  Score 14;  DB 3;  Length 24;
  Best Local Similarity   77.3%;  Pred. No. 9.1e+03;
  Matches   17;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy         23 AACCGGTCGCGGTTATTAAGAA 44
              ||| |||| | || ||||| ||
Db          2 AACTGGTCACCGTGATTAAAAA 23


RESULT 33
US-09-409-670-17
; Sequence 17, Application US/09409670
; Patent No. 6133028
;  GENERAL INFORMATION:
;    APPLICANT:  IMLER, Jean-Luc
;    APPLICANT:  MEHTALI, Majid
;    APPLICANT:  PAVIRANI, Andrea
;    TITLE OF INVENTION:  DEFECTIVE ADENOVIRUSES AND CORRESPONDING
;    TITLE OF INVENTION:  COMPLEMENTATION LINES
;    NUMBER OF SEQUENCES:  43
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  BURNS, DOANE, SWECKER & MATHIS, L.L.P.
;      STREET:  1737 King Street, Suite 500
;      CITY:  Alexandria
;      STATE:  Virginia
;      COUNTRY:  United States
;      ZIP:  22314-2756
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/409,670
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/379,452
;      FILING DATE:  26-JAN-1995
;      APPLICATION NUMBER:  WO PCT/FR94/00624
;      FILING DATE:  27-MAY-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  FR 93 06482
;      FILING DATE:  28-MAY-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Dadio, Susan M.
;      REGISTRATION NUMBER:  40,373
;      REFERENCE/DOCKET NUMBER:  029395-002
;  INFORMATION FOR SEQ ID NO:  17:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  24 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    ORIGINAL SOURCE:
;      ORGANISM:  Synthetic oligonucleotide (OTG5482)
US-09-409-670-17

  Query Match             31.8%;  Score 14;  DB 3;  Length 24;
  Best Local Similarity   77.3%;  Pred. No. 9.1e+03;
  Matches   17;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy         23 AACCGGTCGCGGTTATTAAGAA 44
              ||| |||| | || ||||| ||
Db          2 AACTGGTCACCGTGATTAAAAA 23


RESULT 34
US-09-461-697-456
; Sequence 456, Application US/09461697
; Patent No. 6277974
; GENERAL INFORMATION:
;  APPLICANT: COGENT NEUROSCIENCE, Inc.
;  APPLICANT: Lo, Donald C.
;  APPLICANT: Barney, Shawn
;  APPLICANT: Thomas, Mary Beth
;  APPLICANT: Portbury, Stuart D.
;  APPLICANT: Puranam, Kasturi
;  APPLICANT: Katz, Lawrence C.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR DIAGNOSING
;  TITLE OF INVENTION: AND TREATING CONDITIONS, DISORDERS, OR DISEASES INVOLVING
;  TITLE OF INVENTION: CELL DEATH
;  FILE REFERENCE: 10001-005-999
;  CURRENT APPLICATION NUMBER: US/09/461,697
;  CURRENT FILING DATE: 1999-12-14
;  NUMBER OF SEQ ID NOS: 466
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 456
;   LENGTH: 30
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-461-697-456

  Query Match             31.8%;  Score 14;  DB 3;  Length 30;
  Best Local Similarity   77.3%;  Pred. No. 9.6e+03;
  Matches   17;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy         20 AATAACCGGTCGCGGTTATTAA 41
              || |||||||   |||| ||||
Db          9 AACAACCGGTTTGGGTTGTTAA 30


RESULT 35
US-09-709-103-34
; Sequence 34, Application US/09709103
; Patent No. 6733991
; GENERAL INFORMATION:
;  APPLICANT: Cismowski, Mary
```

```
;  APPLICANT:  Duzic, Emir
;  TITLE OF INVENTION: AGS Proteins and Nucleic Acid Molecules and Uses Therefor
;  FILE REFERENCE: 60388-A-PCT-US
;  CURRENT APPLICATION NUMBER: US/09/709,103
;  CURRENT FILING DATE:  2000-11-08
;  NUMBER OF SEQ ID NOS: 73
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 34
;   LENGTH: 32
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Probe/Primer
US-09-709-103-34

  Query Match             31.8%;  Score 14;  DB 4;  Length 32;
  Best Local Similarity   77.3%;  Pred. No. 9.8e+03;
  Matches   17;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy         23 AACCGGTCGCGGTTATTAAGAA 44
              ||| || |||| | || |||||
Db         10 AACTGGCCGCGATGATCAAGAA 31


RESULT 36
US-09-439-410A-34
; Sequence 34, Application US/09439410A
; Patent No. 6746852
; GENERAL INFORMATION:
;  APPLICANT: Cismowski, Mary
;  APPLICANT:  Duzic, Emir
;  TITLE OF INVENTION: AGS PROTEIN AND NUCLETIC ACID MOLECULES AND USES THEREOF
;  FILE REFERENCE: 1919/60388-B
;  CURRENT APPLICATION NUMBER: US/09/439,410A
;  CURRENT FILING DATE:  1999-11-11
;  NUMBER OF SEQ ID NOS: 118
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 34
;   LENGTH: 32
;   TYPE: DNA
;   ORGANISM: Artificial sequence
;   FEATURE:
;   OTHER INFORMATION: Probe/Primer
US-09-439-410A-34

  Query Match             31.8%;  Score 14;  DB 4;  Length 32;
  Best Local Similarity   77.3%;  Pred. No. 9.8e+03;
  Matches   17;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy         23 AACCGGTCGCGGTTATTAAGAA 44
              ||| || |||| | || |||||
Db         10 AACTGGCCGCGATGATCAAGAA 31


RESULT 37
US-09-709-103-8
; Sequence 8, Application US/09709103
; Patent No. 6733991
; GENERAL INFORMATION:
;  APPLICANT: Cismowski, Mary
;  APPLICANT:  Duzic, Emir
;  TITLE OF INVENTION: AGS Proteins and Nucleic Acid Molecules and Uses Therefor
;  FILE REFERENCE: 60388-A-PCT-US
;  CURRENT APPLICATION NUMBER: US/09/709,103
;  CURRENT FILING DATE:  2000-11-08
;  NUMBER OF SEQ ID NOS: 73
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 8
;   LENGTH: 35
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Probe/Primer
US-09-709-103-8

  Query Match             31.8%;  Score 14;  DB 4;  Length 35;
  Best Local Similarity   77.3%;  Pred. No. 1e+04;
  Matches   17;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy         23 AACCGGTCGCGGTTATTAAGAA 44
              ||| || |||| | || |||||
Db         13 AACTGGCCGCGATGATCAAGAA 34


RESULT 38
US-09-439-410A-8
; Sequence 8, Application US/09439410A
; Patent No. 6746852
; GENERAL INFORMATION:
;  APPLICANT: Cismowski, Mary
;  APPLICANT:  Duzic, Emir
;  TITLE OF INVENTION: AGS PROTEIN AND NUCLETIC ACID MOLECULES AND USES THEREOF
;  FILE REFERENCE: 1919/60388-B
;  CURRENT APPLICATION NUMBER: US/09/439,410A
;  CURRENT FILING DATE:  1999-11-11
;  NUMBER OF SEQ ID NOS: 118
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 8
;   LENGTH: 35
;   TYPE: DNA
;   ORGANISM: Artificial sequence
;   FEATURE:
;   OTHER INFORMATION: Probe/Primer
US-09-439-410A-8

  Query Match             31.8%;  Score 14;  DB 4;  Length 35;
  Best Local Similarity   77.3%;  Pred. No. 1e+04;
  Matches   17;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy         23 AACCGGTCGCGGTTATTAAGAA 44
              ||| || |||| | || |||||
Db         13 AACTGGCCGCGATGATCAAGAA 34


RESULT 39
US-09-060-299-117/c
; Sequence 117, Application US/09060299
; Patent No. 6545137
;  GENERAL INFORMATION:
;    APPLICANT:  Todd, John A
;    APPLICANT:  Hess, John W
;    APPLICANT:  Caskey, Charles T
;    APPLICANT:  Cox, Roger D
;    APPLICANT:  Gerhold, David
;    APPLICANT:  Hammond, Holly
;    APPLICANT:  Hey, Patricia
;    APPLICANT:  Kawaguchi, Yoshihiko
;    APPLICANT:  Merriman, Tony R
;    APPLICANT:  Metzker, Michael L
;    TITLE OF INVENTION:  No. 6545137el Receptor
;    NUMBER OF SEQUENCES:  455
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Nixon and Vanderhye
;      STREET:  1100 No. 6545137th Glebe Road, Eighth Floor
;      CITY:  Arlington
;      STATE:  Virginia
;      COUNTRY:  US
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25 (EPO)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/060,299
;      FILING DATE:  15-APR-1998
;      CLASSIFICATION:  435
```

```
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/043,553
;      FILING DATE:  15-APR-1997
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/048,740
;      FILING DATE:  05-JUN-1997
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  B.J.Sadoff
;      REGISTRATION NUMBER:  36,663
;      REFERENCE/DOCKET NUMBER:  620-35
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (703)816-4091
;      TELEFAX:  (703)816-4100
;  INFORMATION FOR SEQ ID NO:  117:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  41 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-09-060-299-117

  Query Match             31.8%;  Score 14;  DB 4;  Length 41;
  Best Local Similarity   66.7%;  Pred. No. 1e+04;
  Matches   20;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy          9 CGTTCCTTCTTAATAACCGGTCGCGGTTAT 38
              ||||||||||||||     || ||   | ||
Db         39 CGTTCCTTCTTAACTCAGGGGCGTAATCAT 10


RESULT 40
US-09-402-923A-117/c
; Sequence 117, Application US/09402923A
; Patent No. 6555654
;  GENERAL INFORMATION:
;       APPLICANT: Todd, John A
;                  Hess, John W
;                  Caskey, Charles T
;                  Cox, Roger D
;                  Gerhold, David
;                  Hammond, Holly
;                  Hey, Patricia
;                  Kawaguchi, Yoshihiko
;                  Merriman, Tony R
;                  Metzker, Michael L
;       TITLE OF INVENTION: No. 6555654el LDL-Receptor
;       NUMBER OF SEQUENCES: 455
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: Nixon and Vanderhye
;            STREET: 1100 No. 6555654th Glebe Road, Eighth Floor
;            CITY: Arlington
;            STATE: Virginia
;            COUNTRY: US
;            ZIP: VA 22201-4714
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Floppy disk
;            COMPUTER: IBM PC compatible
;            OPERATING SYSTEM: PC-DOS/MS-DOS
;            SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/09/402,923A
;            FILING DATE: 14-Feb-2001
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: PCT/GB98/01102
;            FILING DATE: 15-APR-1998
;            APPLICATION NUMBER: US 60/043,553
;            FILING DATE: 15-APR-1997
;            APPLICATION NUMBER: US 60/048,740
;            FILING DATE: 05-JUN-1997
;       ATTORNEY/AGENT INFORMATION:
;            NAME: B.J.Sadoff
;            REGISTRATION NUMBER: 36,663
;            REFERENCE/DOCKET NUMBER: 620-81
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: (703)816-4091
;            TELEFAX: (703)816-4100
;  INFORMATION FOR SEQ ID NO: 117:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 41 base pairs
;            TYPE: nucleic acid
;            STRANDEDNESS: single
;            TOPOLOGY: linear
;       SEQUENCE DESCRIPTION: SEQ ID NO: 117:
US-09-402-923A-117

  Query Match             31.8%;  Score 14;  DB 4;  Length 41;
  Best Local Similarity   66.7%;  Pred. No. 1e+04;
  Matches   20;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy          9 CGTTCCTTCTTAATAACCGGTCGCGGTTAT 38
              ||||||||||||||     || ||   | ||
Db         39 CGTTCCTTCTTAACTCAGGGGCGTAATCAT 10


Search completed: May 24, 2005, 12:45:35
Job time : 101 secs
```

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: May 24, 2005, 12:10:04 ; Search time 339 Seconds
(without alignments)
795.936 Million cell updates/sec

Title: US-10-673-063-3_COPY_900_943
Perfect score: 44
Sequence: 1 gcgggtcccgttccttctta.....ccggtcgcggttattaagaa 44

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 5695437 seqs, 3066160638 residues

Total number of hits satisfying chosen parameters: 5377818

Minimum DB seq length: 0
Maximum DB seq length: 50

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 100 summaries

Database : Published_Applications_NA:*
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
16: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:*
17: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:*
18: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq:*
19: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
20: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq:*
21: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
22: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
c	1	16	36.4	27	9	US-09-854-122-8	Sequence 8, Appli
	2	16	36.4	27	9	US-09-854-122-9	Sequence 9, Appli
	3	16	36.4	30	19	US-10-823-259-45	Sequence 45, Appl
	4	16	36.4	30	19	US-10-823-254-45	Sequence 45, Appl
c	5	16	36.4	33	19	US-10-792-498-20	Sequence 20, Appl
c	6	15.8	35.9	25	19	US-10-809-189-29372	Sequence 29372, A
	7	15.8	35.9	41	17	US-10-035-833A-1851	Sequence 1851, Ap
	8	15.8	35.9	41	17	US-10-035-833A-4446	Sequence 4446, Ap
c	9	15.6	35.5	33	18	US-10-466-347-11	Sequence 11, Appl
c	10	15.4	35.0	25	18	US-10-719-895-20	Sequence 20, Appl
c	11	15.4	35.0	25	19	US-10-719-900-918103	Sequence 918103,
c	12	15.2	34.5	47	18	US-10-343-561-98	Sequence 98, Appl
c	13	15	34.1	25	19	US-10-719-900-350576	Sequence 350576,
	14	15	34.1	25	19	US-10-719-900-825969	Sequence 825969,
c	15	15	34.1	31	9	US-09-801-274-523	Sequence 523, App
	16	15	34.1	50	17	US-10-131-827-2605	Sequence 2605, Ap
c	17	15	34.1	50	17	US-10-131-827-7716	Sequence 7716, Ap
	18	14.8	33.6	25	19	US-10-719-900-81188	Sequence 81188, A
c	19	14.8	33.6	25	19	US-10-719-900-927327	Sequence 927327,
c	20	14.8	33.6	25	19	US-10-809-189-29373	Sequence 29373, A
c	21	14.8	33.6	31	10	US-09-927-046-4154	Sequence 4154, Ap
c	22	14.8	33.6	47	17	US-10-349-143-240	Sequence 240, App
c	23	14.8	33.6	47	17	US-10-349-143-1725	Sequence 1725, Ap
c	24	14.8	33.6	50	10	US-09-993-346-585	Sequence 585, App
	25	14.8	33.6	50	17	US-10-131-827-2809	Sequence 2809, Ap
	26	14.6	33.2	25	15	US-10-098-263B-97505	Sequence 97505, A
	27	14.6	33.2	25	19	US-10-719-900-202404	Sequence 202404,
	28	14.6	33.2	25	19	US-10-719-900-728815	Sequence 728815,
	29	14.6	33.2	25	19	US-10-719-900-797256	Sequence 797256,
	30	14.6	33.2	25	19	US-10-719-900-816976	Sequence 816976,
	31	14.6	33.2	25	19	US-10-719-900-854921	Sequence 854921,
	32	14.6	33.2	25	19	US-10-719-900-896407	Sequence 896407,
c	33	14.6	33.2	35	19	US-10-489-739-7	Sequence 7, Appli
c	34	14.6	33.2	41	17	US-10-252-155-747	Sequence 747, App
c	35	14.6	33.2	45	9	US-09-991-003B-16	Sequence 16, Appl
	36	14.6	33.2	47	17	US-10-349-143-3415	Sequence 3415, Ap
c	37	14.4	32.7	25	15	US-10-098-263B-88885	Sequence 88885, A
	38	14.4	32.7	25	19	US-10-719-900-221631	Sequence 221631,
	39	14.4	32.7	25	19	US-10-719-900-221632	Sequence 221632,
c	40	14.4	32.7	25	19	US-10-719-900-434524	Sequence 434524,
	41	14.4	32.7	25	19	US-10-719-900-639025	Sequence 639025,
	42	14.4	32.7	34	14	US-10-025-222A-7	Sequence 7, Appli
	43	14.4	32.7	41	17	US-10-252-155-747	Sequence 747, App
	44	14.4	32.7	47	17	US-10-349-143-2970	Sequence 2970, Ap
c	45	14.4	32.7	50	17	US-10-062-188-128	Sequence 128, App
c	46	14.4	32.7	50	17	US-10-062-188-159	Sequence 159, App
c	47	14.2	32.3	20	10	US-09-922-146-31	Sequence 31, Appl
c	48	14.2	32.3	20	17	US-10-289-762-6719	Sequence 6719, Ap
c	49	14.2	32.3	21	18	US-10-786-720-7147	Sequence 7147, Ap
c	50	14.2	32.3	21	18	US-10-786-720-9415	Sequence 9415, Ap
c	51	14.2	32.3	21	19	US-10-848-755A-25	Sequence 25, Appl
c	52	14.2	32.3	32	17	US-10-371-771-16	Sequence 16, Appl
	53	14.2	32.3	33	18	US-10-799-372-2	Sequence 2, Appli
	54	14.2	32.3	39	17	US-10-423-828-54	Sequence 54, Appl
c	55	14.2	32.3	50	17	US-10-175-689-10	Sequence 10, Appl
c	56	14.2	32.3	50	17	US-10-131-827-4778	Sequence 4778, Ap
	57	14	31.8	24	9	US-09-725-720-17	Sequence 17, Appl
	58	14	31.8	24	10	US-09-739-007-17	Sequence 17, Appl
	59	14	31.8	24	17	US-10-169-027-1	Sequence 1, Appli
	60	14	31.8	25	19	US-10-719-900-118945	Sequence 118945,
	61	14	31.8	25	19	US-10-719-900-118965	Sequence 118965,
	62	14	31.8	25	19	US-10-719-900-328268	Sequence 328268,
	63	14	31.8	25	19	US-10-719-900-328269	Sequence 328269,
	64	14	31.8	25	19	US-10-719-900-349040	Sequence 349040,
	65	14	31.8	25	19	US-10-719-900-349041	Sequence 349041,
	66	14	31.8	25	19	US-10-719-900-350580	Sequence 350580,
c	67	14	31.8	25	19	US-10-719-900-572728	Sequence 572728,
	68	14	31.8	25	19	US-10-719-900-804828	Sequence 804828,
c	69	14	31.8	25	19	US-10-719-900-818986	Sequence 818986,
	70	14	31.8	30	9	US-09-922-261-456	Sequence 456, App
	71	14	31.8	32	18	US-10-804-491-34	Sequence 34, Appl
	72	14	31.8	35	18	US-10-804-491-8	Sequence 8, Appli
	73	14	31.8	38	15	US-10-282-287-29	Sequence 29, Appl
c	74	14	31.8	41	16	US-10-331-907-117	Sequence 117, App
c	75	14	31.8	42	18	US-10-804-408-96	Sequence 96, Appl
	76	14	31.8	45	14	US-10-294-171-5	Sequence 5, Appli
	77	14	31.8	47	10	US-09-993-346-250	Sequence 250, App
	78	14	31.8	47	17	US-10-349-143-990	Sequence 990, App
c	79	13.8	31.4	21	18	US-10-786-720-7148	Sequence 7148, Ap
	80	13.8	31.4	21	18	US-10-786-720-7149	Sequence 7149, Ap
c	81	13.8	31.4	21	18	US-10-786-720-9416	Sequence 9416, Ap
	82	13.8	31.4	21	18	US-10-786-720-9417	Sequence 9417, Ap
c	83	13.8	31.4	24	10	US-09-940-185-1982	Sequence 1982, Ap
	84	13.8	31.4	25	14	US-10-215-112-12954	Sequence 12954, A

c	85	13.8	31.4	25	15	US-10-098-263B-71257	Sequence 71257, A
c	86	13.8	31.4	25	15	US-10-098-263B-71258	Sequence 71258, A
	87	13.8	31.4	25	15	US-10-098-263B-87670	Sequence 87670, A
c	88	13.8	31.4	25	15	US-10-098-263B-124443	Sequence 124443,
	89	13.8	31.4	25	15	US-10-098-263B-130243	Sequence 130243,
	90	13.8	31.4	25	19	US-10-719-900-389023	Sequence 389023,
	91	13.8	31.4	25	19	US-10-719-900-389024	Sequence 389024,
	92	13.8	31.4	25	19	US-10-719-900-690492	Sequence 690492,
c	93	13.8	31.4	25	19	US-10-719-900-918104	Sequence 918104,
	94	13.8	31.4	25	19	US-10-809-189-50594	Sequence 50594, A
	95	13.8	31.4	26	14	US-10-224-836-190	Sequence 190, App
c	96	13.8	31.4	26	16	US-10-347-183-11	Sequence 11, Appl
c	97	13.8	31.4	38	10	US-09-861-257-99	Sequence 99, Appl
c	98	13.8	31.4	38	15	US-10-189-360-124	Sequence 124, App
c	99	13.8	31.4	47	9	US-09-908-599-20	Sequence 20, Appl
c	100	13.8	31.4	47	9	US-09-908-599-22	Sequence 22, Appl

ALIGNMENTS

```
RESULT 1
US-09-854-122-8/c
; Sequence 8, Application US/09854122
; Patent No. US20020016980A1
; GENERAL INFORMATION:
;  APPLICANT: ALBERTE, RANDALL S.
;  APPLICANT:  SMITH, ROBERT
;  TITLE OF INVENTION: TRANSGENIC PLANTS INCORPORATING TRAIT OF ZOSTERA MARINA
;  FILE REFERENCE: PHA-007.01
;  CURRENT APPLICATION NUMBER: US/09/854,122
;  CURRENT FILING DATE:  2001-09-10
;  PRIOR APPLICATION NUMBER: 60/202,529
;  PRIOR FILING DATE: 2000-05-10
;  NUMBER OF SEQ ID NOS: 51
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 8
;   LENGTH: 27
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: Primer
US-09-854-122-8

  Query Match             36.4%;  Score 16;  DB 9;  Length 27;
  Best Local Similarity   79.2%;  Pred. No. 5.8e+03;
  Matches   19;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy         20 AATAACCGGTCGCGGTTATTAAGA 43
              ||||||  |||| ||||| | |||
Db         25 AATAACTTGTCGGGGTTAATCAGA 2


RESULT 2
US-09-854-122-9
; Sequence 9, Application US/09854122
; Patent No. US20020016980A1
; GENERAL INFORMATION:
;  APPLICANT: ALBERTE, RANDALL S.
;  APPLICANT:  SMITH, ROBERT
;  TITLE OF INVENTION: TRANSGENIC PLANTS INCORPORATING TRAIT OF ZOSTERA MARINA
;  FILE REFERENCE: PHA-007.01
;  CURRENT APPLICATION NUMBER: US/09/854,122
;  CURRENT FILING DATE:  2001-09-10
;  PRIOR APPLICATION NUMBER: 60/202,529
;  PRIOR FILING DATE: 2000-05-10
;  NUMBER OF SEQ ID NOS: 51
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
;   LENGTH: 27
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: Primer
US-09-854-122-9

  Query Match             36.4%;  Score 16;  DB 9;  Length 27;
  Best Local Similarity   79.2%;  Pred. No. 5.8e+03;
  Matches   19;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy         20 AATAACCGGTCGCGGTTATTAAGA 43
              ||||||  |||| ||||| | |||
Db          3 AATAACTTGTCGGGGTTAATCAGA 26


RESULT 3
US-10-823-259-45
; Sequence 45, Application US/10823259
; Publication No. US20050049176A1
; GENERAL INFORMATION:
;  APPLICANT: Kiener, Peter
;  APPLICANT:  Kinch, Michael
;  APPLICANT:  Langermann, Solomon
;  TITLE OF INVENTION: EphA2 and Hyperproliferative Cell Disorders and Epithelial and
;  TITLE OF INVENTION:  Endothelial Reconstitution
;  FILE REFERENCE: 10271-058-999
;  CURRENT APPLICATION NUMBER: US/10/823,259
;  CURRENT FILING DATE:  2004-04-12
;  PRIOR APPLICATION NUMBER: 60/462,009
;  PRIOR FILING DATE: 2003-04-11
;  NUMBER OF SEQ ID NOS: 45
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 45
;   LENGTH: 30
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of artificial sequence:  phosphorothioate-modified ant
;   OTHER INFORMATION: oligonucleotides
US-10-823-259-45

  Query Match             36.4%;  Score 16;  DB 19;  Length 30;
  Best Local Similarity   79.2%;  Pred. No. 6e+03;
  Matches   19;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          2 CGGGTCCCGTTCCTTCTTAATAAC 25
              || |||||||||||||   || ||
Db          3 CGCGTCCCGTTCCTTCACCATGAC 26


RESULT 4
US-10-823-254-45
; Sequence 45, Application US/10823254
; Publication No. US20050059592A1
; GENERAL INFORMATION:
;  APPLICANT: Kiener, Peter
;  APPLICANT:  Kinch, Michael
;  APPLICANT:  Langermann, Solomon
;  APPLICANT:  Reed, Jennifer
;  TITLE OF INVENTION: EphA2 and Hyperproliferative Cell Disorders
;  FILE REFERENCE: 10271-060-999
;  CURRENT APPLICATION NUMBER: US/10/823,254
;  CURRENT FILING DATE:  2004-04-12
;  PRIOR APPLICATION NUMBER: 60/462,024
;  PRIOR FILING DATE: 2003-04-11
;  NUMBER OF SEQ ID NOS: 45
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 45
;   LENGTH: 30
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of artificial sequence:  phosphorothioate-modified an
US-10-823-254-45

  Query Match             36.4%;  Score 16;  DB 19;  Length 30;
```

Best Local Similarity 79.2%; Pred. No. 6e+03;
Matches 19; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

```
Qy          2 CGGGTCCCGTTCCTTCTTAATAAC 25
              || |||||||||||||   || ||
Db          3 CGCGTCCCGTTCCTTCACCATGAC 26
```

RESULT 5
US-10-792-498-20/c
; Sequence 20, Application US/10792498
; Publication No. US20050074865A1
; GENERAL INFORMATION:
; APPLICANT: Afeyan, Noubar B.
; APPLICANT: Lee, Frank D.
; APPLICANT: Wong, Gordon G.
; APPLICANT: Das Gupta, Ruchira
; APPLICANT: Baynes, Brian
; TITLE OF INVENTION: ADZYMES AND USES THEREOF
; FILE REFERENCE: COTH-P03-001
; CURRENT APPLICATION NUMBER: US/10/792,498
; CURRENT FILING DATE: 2004-03-02
; PRIOR APPLICATION NUMBER: US 10/650,592
; PRIOR FILING DATE: 2003-08-27
; PRIOR APPLICATION NUMBER: US 60/406,517
; PRIOR FILING DATE: 2002-08-27
; PRIOR APPLICATION NUMBER: US 60/423,754
; PRIOR FILING DATE: 2002-11-05
; PRIOR APPLICATION NUMBER: US 60/430,001
; PRIOR FILING DATE: 2002-11-27
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 20
; LENGTH: 33
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Oligo prethrombinfwdH3
US-10-792-498-20

Query Match 36.4%; Score 16; DB 19; Length 33;
Best Local Similarity 79.2%; Pred. No. 6.1e+03;
Matches 19; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

```
Qy         19 TAATAACCGGTCGCGGTTATTAAG 42
              || | || ||| ||||| ||||||
Db         32 TACTCACTGGTGGCGGTCATTAAG 9
```

RESULT 6
US-10-809-189-29372/c
; Sequence 29372, Application US/10809189
; Publication No. US20050048531A1
; GENERAL INFORMATION:
; APPLICANT: Michael Mittmann
; APPLICANT: David Mack
; APPLICANT: David Lockhart
; APPLICANT: Affymetrix, Inc.
; TITLE OF INVENTION: Methods of Genetic Analysis
; FILE REFERENCE: 3101.1
; CURRENT APPLICATION NUMBER: US/10/809,189
; CURRENT FILING DATE: 2004-03-25
; PRIOR APPLICATION NUMBER: US/09/396,196
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: 60/100,678
; PRIOR FILING DATE: 1998-09-17
; NUMBER OF SEQ ID NOS: 127806
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 29372
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Mus musculus
US-10-809-189-29372

Query Match 35.9%; Score 15.8; DB 19; Length 25;
Best Local Similarity 89.5%; Pred. No. 7e+03;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy         26 CGGTCGCGGTTATTAAGAA 44
              | |||||||| ||||||||
Db         19 CAGTCGCGGTAATTAAGAA 1
```

RESULT 7
US-10-035-833A-1851
; Sequence 1851, Application US/10035833A
; Publication No. US20040072156A1
; GENERAL INFORMATION:
; APPLICANT: Nakamura, Yuho
; APPLICANT: Sekine, Akihiro
; APPLICANT: Iida, Aritoshi
; APPLICANT: Saito, Osamu
; TITLE OF INVENTION: Detection of Genetic Polymorphisms
; FILE REFERENCE: FORS-06904
; CURRENT APPLICATION NUMBER: US/10/035,833A
; CURRENT FILING DATE: 2001-12-27
; NUMBER OF SEQ ID NOS: 7669
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1851
; LENGTH: 41
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-035-833A-1851

Query Match 35.9%; Score 15.8; DB 17; Length 41;
Best Local Similarity 81.0%; Pred. No. 7.9e+03;
Matches 17; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

```
Qy          5 GTCCCGTTCCTTCTTAATAAC 25
              |||||:||||||| | ||| |
Db         16 GTCCCRTTCCTTCATTATATC 36
```

RESULT 8
US-10-035-833A-4446
; Sequence 4446, Application US/10035833A
; Publication No. US20040072156A1
; GENERAL INFORMATION:
; APPLICANT: Nakamura, Yuho
; APPLICANT: Sekine, Akihiro
; APPLICANT: Iida, Aritoshi
; APPLICANT: Saito, Osamu
; TITLE OF INVENTION: Detection of Genetic Polymorphisms
; FILE REFERENCE: FORS-06904
; CURRENT APPLICATION NUMBER: US/10/035,833A
; CURRENT FILING DATE: 2001-12-27
; NUMBER OF SEQ ID NOS: 7669
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 4446
; LENGTH: 41
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-035-833A-4446

Query Match 35.9%; Score 15.8; DB 17; Length 41;
Best Local Similarity 81.0%; Pred. No. 7.9e+03;
Matches 17; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

```
Qy          5 GTCCCGTTCCTTCTTAATAAC 25
              |||||:||||||| | ||| |
Db         16 GTCCCRTTCCTTCATTATATC 36
```

RESULT 9
US-10-466-347-11/c
; Sequence 11, Application US/10466347

```
; Publication No. US20040109849A1
; GENERAL INFORMATION:
;  APPLICANT: Fazio, Vito M.
;  TITLE OF INVENTION: DNA VACCINES EXPRESSING HYPERVARIABLE VH-CDR3 IDIOTYPIC DETERMINA
;  FILE REFERENCE: 02901/000N028-US0
;  CURRENT APPLICATION NUMBER: US/10/466,347
;  CURRENT FILING DATE:  2003-12-08
;  PRIOR APPLICATION NUMBER: PCT/IT01/00014
;  PRIOR FILING DATE: 2002-07-18
;  NUMBER OF SEQ ID NOS: 17
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 11
;   LENGTH: 33
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: PCR primer
US-10-466-347-11

  Query Match             35.5%;  Score 15.6;  DB 18;  Length 33;
  Best Local Similarity   70.0%;  Pred. No. 9.1e+03;
  Matches   21;  Conservative    0;  Mismatches    9;  Indels    0;  Gaps    0;

Qy          3 GGGTCCCGTTCCTTCTTAATAACCGGTCGC 32
              ||   ||||  | || |||||| ||| |||
Db         33 GGTCACCGTCTCCTCATAATAAGCGGCCGC 4


RESULT 10
US-10-719-895-20/c
; Sequence 20, Application US/10719895
; Publication No. US20040213805A1
; GENERAL INFORMATION:
;  APPLICANT: Verheije, Monique H.
;  TITLE OF INVENTION: Deletions in Arterivirus replicons
;  FILE REFERENCE: P55434US
;  CURRENT APPLICATION NUMBER: US/10/719,895
;  CURRENT FILING DATE:  2003-11-21
;  PRIOR APPLICATION NUMBER: EP 01201921.2
;  PRIOR FILING DATE: 2001-05-21
;  NUMBER OF SEQ ID NOS: 52
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 20
;   LENGTH: 25
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Primer LV266
US-10-719-895-20

  Query Match             35.0%;  Score 15.4;  DB 18;  Length 25;
  Best Local Similarity   76.0%;  Pred. No. 1e+04;
  Matches   19;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          6 TCCCGTTCCTTCTTAATAACCGGTC 30
              || |||  ||||||||| | | |||
Db         25 TCGCGTGACTTCTTAATTAACAGTC 1


RESULT 11
US-10-719-900-918103/c
; Sequence 918103, Application US/10719900
; Publication No. US20050026164A1
; GENERAL INFORMATION:
;  APPLICANT: Xue Mei Zhou
;  TITLE OF INVENTION: Methods of Genetic Analysis of Mouse
;  FILE REFERENCE: 3528.1
;  CURRENT APPLICATION NUMBER: US/10/719,900
;  CURRENT FILING DATE:  2003-11-20
;  PRIOR APPLICATION NUMBER: 60/427,808
;  PRIOR FILING DATE: 2002 11 20
;  NUMBER OF SEQ ID NOS: 982914
;  SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
; SEQ ID NO 918103
;   LENGTH: 25
;   TYPE: DNA
;   ORGANISM: Mus musculus
US-10-719-900-918103

  Query Match             35.0%;  Score 15.4;  DB 19;  Length 25;
  Best Local Similarity   76.0%;  Pred. No. 1e+04;
  Matches   19;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy         17 CTTAATAACCGGTCGCGGTTATTAA 41
              || || || | || || ||||||||
Db         25 CTGAAGAATCTGTAGCTGTTATTAA 1


RESULT 12
US-10-343-561-98/c
; Sequence 98, Application US/10343561
; Publication No. US20040126389A1
; GENERAL INFORMATION:
;  APPLICANT: Berthet, Francois-Xavier Jacques
;  APPLICANT:  Dalemans, Wilfried
;  APPLICANT:  Denoel, Philippe
;  APPLICANT:  Dequesne, Guy
;  APPLICANT:  Feron, Christiane
;  APPLICANT:  Garcon, Nathalie
;  APPLICANT:  Lobet, Yves
;  APPLICANT:  Poolman, Jan
;  APPLICANT:  Thiry, Georges
;  APPLICANT:  Thonnard, Joelle
;  APPLICANT:  Voet, Pierre
;  TITLE OF INVENTION: Vaccines Comprising Outer Membrane
;  TITLE OF INVENTION:  Vesicles from Gram Negative Bacteria
;  FILE REFERENCE: B45260
;  CURRENT APPLICATION NUMBER: US/10/343,561
;  CURRENT FILING DATE:  2003-01-31
;  PRIOR APPLICATION NUMBER: PCT/EP01/08857
;  PRIOR FILING DATE: 2001-07-31
;  PRIOR APPLICATION NUMBER: EP 00956369.3
;  PRIOR FILING DATE: 2000-07-31
;  PRIOR APPLICATION NUMBER: GB 0103170.7
;  PRIOR FILING DATE: 2001-02-08
;  NUMBER OF SEQ ID NOS: 156
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 98
;   LENGTH: 47
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: PNS4 primer
US-10-343-561-98

  Query Match             34.5%;  Score 15.2;  DB 18;  Length 47;
  Best Local Similarity   71.4%;  Pred. No. 1.5e+04;
  Matches   20;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Qy         17 CTTAATAACCGGTCGCGGTTATTAAGAA 44
              | || |  |||| |||| | ||||||
Db         41 CATATTTTCCGGACGCGTTAATTAAGGA 14


RESULT 13
US-10-719-900-350576/c
; Sequence 350576, Application US/10719900
; Publication No. US20050026164A1
; GENERAL INFORMATION:
;  APPLICANT: Xue Mei Zhou
;  TITLE OF INVENTION: Methods of Genetic Analysis of Mouse
;  FILE REFERENCE: 3528.1
;  CURRENT APPLICATION NUMBER: US/10/719,900
;  CURRENT FILING DATE:  2003-11-20
;  PRIOR APPLICATION NUMBER: 60/427,808
;  PRIOR FILING DATE: 2002 11 20
```

```
;  NUMBER OF SEQ ID NOS: 982914
;  SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
; SEQ ID NO 350576
;    LENGTH: 25
;    TYPE: DNA
;    ORGANISM: Mus musculus
US-10-719-900-350576

  Query Match             34.1%;  Score 15;  DB 19;  Length 25;
  Best Local Similarity   78.3%;  Pred. No. 1.5e+04;
  Matches   18;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy         12 TCCTTCTTAATAACCGGTCGCGG 34
              ||   ||||||||| |||| |||
Db         23 TCTGCCTTAATAACTGGTCCCGG 1


RESULT 14
US-10-719-900-825969
; Sequence 825969, Application US/10719900
; Publication No. US20050026164A1
; GENERAL INFORMATION:
;  APPLICANT: Xue Mei Zhou
;  TITLE OF INVENTION: Methods of Genetic Analysis of Mouse
;  FILE REFERENCE: 3528.1
;  CURRENT APPLICATION NUMBER: US/10/719,900
;  CURRENT FILING DATE:  2003-11-20
;  PRIOR APPLICATION NUMBER: 60/427,808
;  PRIOR FILING DATE: 2002 11 20
;  NUMBER OF SEQ ID NOS: 982914
;  SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
; SEQ ID NO 825969
;    LENGTH: 25
;    TYPE: DNA
;    ORGANISM: Mus musculus
US-10-719-900-825969

  Query Match             34.1%;  Score 15;  DB 19;  Length 25;
  Best Local Similarity   78.3%;  Pred. No. 1.5e+04;
  Matches   18;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          1 GCGGGTCCCGTTCCTTCTTAATA 23
              | || ||| ||| | |||||||||
Db          3 GTGGTTCCAGTTGCATCTTAATA 25


RESULT 15
US-09-801-274-523/c
; Sequence 523, Application US/09801274
; Patent No. US20020032319A1
; GENERAL INFORMATION:
;  APPLICANT: Cargill, Michele
;  APPLICANT:  Ireland, James S.
;  APPLICANT:  Lander, Eric S.
;  TITLE OF INVENTION: HUMAN SINGLE NUCLEOTIDE POLYMORPHISMS
;  FILE REFERENCE: 2825.2009-001
;  CURRENT APPLICATION NUMBER: US/09/801,274
;  CURRENT FILING DATE:  2001-03-07
;  PRIOR APPLICATION NUMBER: US 60/187,510
;  PRIOR FILING DATE: 2000-03-07
;  PRIOR APPLICATION NUMBER: US 60/206,129
;  PRIOR FILING DATE: 2000-05-22
;  NUMBER OF SEQ ID NOS: 1802
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 523
;    LENGTH: 31
;    TYPE: DNA
;    ORGANISM: Homo sapiens
US-09-801-274-523

  Query Match             34.1%;  Score 15;  DB 9;  Length 31;
  Best Local Similarity   78.3%;  Pred. No. 1.6e+04;
  Matches   18;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy         12 TCCTTCTTAATAACCGGTCGCGG 34
              || |||||||| | | |||| ||
Db         29 TCTTTCTTAATGAKCTGTCGGGG 7


RESULT 16
US-10-131-827-2605
; Sequence 2605, Application US/10131827
; Publication No. US20040009479A1
; GENERAL INFORMATION:
;  APPLICANT: Wohlgemuth, Jay
;  APPLICANT:  Fry, Kirk
;  APPLICANT:  Woodward, Robert
;  APPLICANT:  Ly, Ngoc
;  TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING AND MONITORING AUTOIMMUNE
;  TITLE OF INVENTION:  CHRONIC INFLAMMATORY DISEASES
;  FILE REFERENCE: 506612000120
;  CURRENT APPLICATION NUMBER: US/10/131,827
;  CURRENT FILING DATE:  2002-09-06
;  PRIOR APPLICATION NUMBER: US 10/006,290
;  PRIOR FILING DATE: 2001-10-22
;  PRIOR APPLICATION NUMBER: US 60/296,764
;  PRIOR FILING DATE: 2001-06-08
;  NUMBER OF SEQ ID NOS: 9090
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2605
;    LENGTH: 50
;    TYPE: DNA
;    ORGANISM: Homo sapiens
US-10-131-827-2605

  Query Match             34.1%;  Score 15;  DB 17;  Length 50;
  Best Local Similarity   67.7%;  Pred. No. 1.8e+04;
  Matches   21;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy         14 CTTCTTAATAACCGGTCGCGGTTATTAAGAA 44
              || ||| ||  ||  | ||| |||||  |||
Db          2 CTGCTTCATCTCCTTTTGCGCTTATTTGGAA 32


RESULT 17
US-10-131-827-7716/c
; Sequence 7716, Application US/10131827
; Publication No. US20040009479A1
; GENERAL INFORMATION:
;  APPLICANT: Wohlgemuth, Jay
;  APPLICANT:  Fry, Kirk
;  APPLICANT:  Woodward, Robert
;  APPLICANT:  Ly, Ngoc
;  TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING AND MONITORING AUTOIMMUNE
;  TITLE OF INVENTION:  CHRONIC INFLAMMATORY DISEASES
;  FILE REFERENCE: 506612000120
;  CURRENT APPLICATION NUMBER: US/10/131,827
;  CURRENT FILING DATE:  2002-09-06
;  PRIOR APPLICATION NUMBER: US 10/006,290
;  PRIOR FILING DATE: 2001-10-22
;  PRIOR APPLICATION NUMBER: US 60/296,764
;  PRIOR FILING DATE: 2001-06-08
;  NUMBER OF SEQ ID NOS: 9090
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 7716
;    LENGTH: 50
;    TYPE: DNA
;    ORGANISM: Homo sapiens
US-10-131-827-7716

  Query Match             34.1%;  Score 15;  DB 17;  Length 50;
  Best Local Similarity   67.7%;  Pred. No. 1.8e+04;
  Matches   21;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy         12 TCCTTCTTAATAACCGGTCGCGGTTATTAAG 42
              | |||||| ||||   |||||  ||   |||
```

```
Db         49 TTCTTCTTCATAATGAGTCGCTTTTGAAAAG 19


RESULT 18
US-10-719-900-81188
; Sequence 81188, Application US/10719900
; Publication No. US20050026164A1
; GENERAL INFORMATION:
;  APPLICANT: Xue Mei Zhou
;  TITLE OF INVENTION: Methods of Genetic Analysis of Mouse
;  FILE REFERENCE: 3528.1
;  CURRENT APPLICATION NUMBER: US/10/719,900
;  CURRENT FILING DATE:  2003-11-20
;  PRIOR APPLICATION NUMBER: 60/427,808
;  PRIOR FILING DATE: 2002 11 20
;  NUMBER OF SEQ ID NOS: 982914
;  SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
; SEQ ID NO 81188
;   LENGTH: 25
;   TYPE: DNA
;   ORGANISM: Mus musculus
US-10-719-900-81188

  Query Match             33.6%;  Score 14.8;  DB 19;  Length 25;
  Best Local Similarity   88.9%;  Pred. No. 1.9e+04;
  Matches   16;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          6 TCCCGTTCCTTCTTAATA 23
              ||| ||| ||||||||||
Db          8 TCCTGTTACTTCTTAATA 25


RESULT 19
US-10-719-900-927327/c
; Sequence 927327, Application US/10719900
; Publication No. US20050026164A1
; GENERAL INFORMATION:
;  APPLICANT: Xue Mei Zhou
;  TITLE OF INVENTION: Methods of Genetic Analysis of Mouse
;  FILE REFERENCE: 3528.1
;  CURRENT APPLICATION NUMBER: US/10/719,900
;  CURRENT FILING DATE:  2003-11-20
;  PRIOR APPLICATION NUMBER: 60/427,808
;  PRIOR FILING DATE: 2002 11 20
;  NUMBER OF SEQ ID NOS: 982914
;  SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
; SEQ ID NO 927327
;   LENGTH: 25
;   TYPE: DNA
;   ORGANISM: Mus musculus
US-10-719-900-927327

  Query Match             33.6%;  Score 14.8;  DB 19;  Length 25;
  Best Local Similarity   88.9%;  Pred. No. 1.9e+04;
  Matches   16;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy         24 ACCGGTCGCGGTTATTAA 41
              ||| |||||| |||||||
Db         22 ACCTGTCGCGATTATTAA 5


RESULT 20
US-10-809-189-29373/c
; Sequence 29373, Application US/10809189
; Publication No. US20050048531A1
; GENERAL INFORMATION:
;  APPLICANT: Michael Mittmann
;  APPLICANT:  David Mack
;  APPLICANT:  David Lockhart
;  APPLICANT:  Affymetrix, Inc.
;  TITLE OF INVENTION: Methods of Genetic Analysis
;  FILE REFERENCE: 3101.1
;  CURRENT APPLICATION NUMBER: US/10/809,189
;  CURRENT FILING DATE:  2004-03-25
;  PRIOR APPLICATION NUMBER: US/09/396,196
;  PRIOR FILING DATE: 1999-09-15
;  PRIOR APPLICATION NUMBER: 60/100,678
;  PRIOR FILING DATE: 1998-09-17
;  NUMBER OF SEQ ID NOS: 127806
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 29373
;   LENGTH: 25
;   TYPE: DNA
;   ORGANISM: Mus musculus
US-10-809-189-29373

  Query Match             33.6%;  Score 14.8;  DB 19;  Length 25;
  Best Local Similarity   88.9%;  Pred. No. 1.9e+04;
  Matches   16;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy         26 CGGTCGCGGTTATTAAGA 43
              | |||||||| |||||||
Db         18 CAGTCGCGGTAATTAAGA 1


RESULT 21
US-09-927-046-4154/c
; Sequence 4154, Application US/09927046
; Publication No. US20030064946A1
; GENERAL INFORMATION:
;  APPLICANT: Ribozyme Pharmaceuticals, Inc
;  APPLICANT:  McSwiggen, Jim
;  APPLICANT:  Thompson, Jim
;  APPLICANT:  McKenzie, Tim
;  APPLICANT:  Ayers, Dave
;  APPLICANT:  Grupe, Andrew
;  APPLICANT:  Szymkowski, Edmund
;  TITLE OF INVENTION: Method and Reagent for the Inhibition of Calcium Activated Chloric
;  TITLE OF INVENTION:  Channel-1
;  FILE REFERENCE: 249/021
;  CURRENT APPLICATION NUMBER: US/09/927,046
;  CURRENT FILING DATE:  2001-08-09
;  NUMBER OF SEQ ID NOS: 5450
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4154
;   LENGTH: 31
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:  Enzymatic Nucleic Acid
US-09-927-046-4154

  Query Match             33.6%;  Score 14.8;  DB 10;  Length 31;
  Best Local Similarity   73.1%;  Pred. No. 2e+04;
  Matches   19;  Conservative    0;  Mismatches    7;  Indels    0;  Gaps    0;

Qy          5 GTCCCGTTCCTTCTTAATAACCGGTC 30
              ||||||||| || |    || || |||
Db         30 GTCCCGTTCGTTGTAGCTAGCCCGTC 5


RESULT 22
US-10-349-143-240/c
; Sequence 240, Application US/10349143
; Publication No. US20040005584A1
; GENERAL INFORMATION:
;  APPLICANT: Cohen, Daniel
;  APPLICANT:  Blumenfeld, Marta
;  APPLICANT:  Chumakov, Ilya
;  TITLE OF INVENTION: Biallelic markers for use in constructing a high density...
;  FILE REFERENCE: GENSET.020CP1
;  CURRENT APPLICATION NUMBER: US/10/349,143
;  CURRENT FILING DATE:  2003-01-21
;  PRIOR APPLICATION NUMBER: US/09/422,978
;  PRIOR FILING DATE: 1999-10-20
;  PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 09/298,850
```

```
;  PRIOR FILING DATE: EARLIER FILING DATE: 1999-04-21
;  PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/109,732
;  PRIOR FILING DATE: EARLIER FILING DATE: 1998-11-23
;  PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/082,614
;  PRIOR FILING DATE: EARLIER FILING DATE: 1998-04-21
;  NUMBER OF SEQ ID NOS: 11796
; SEQ ID NO 240
;   LENGTH: 47
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
;   FEATURE:
;   NAME/KEY: allele
;   LOCATION: 24
;   OTHER INFORMATION: 99-1368-299 : polymorphic base C or T
US-10-349-143-240

  Query Match             33.6%;  Score 14.8;  DB 17;  Length 47;
  Best Local Similarity   61.1%;  Pred. No. 2.2e+04;
  Matches   22;  Conservative    1;  Mismatches   13;  Indels    0;  Gaps    0;

Qy          9 CGTTCCTTCTTAATAACCGGTCGCGGTTATTAAGAA 44
              | ||   | ||||||   | ||:| ||| | || ||
Db         46 CATTTAATTTTAATACATGCTCRCTGTTTTGAAAAA 11


RESULT 23
US-10-349-143-1725/c
; Sequence 1725, Application US/10349143
; Publication No. US20040005584A1
; GENERAL INFORMATION:
;  APPLICANT: Cohen, Daniel
;  APPLICANT:  Blumenfeld, Marta
;  APPLICANT:  Chumakov, Ilya
;  TITLE OF INVENTION: Biallelic markers for use in constructing a high density...
;  FILE REFERENCE: GENSET.020CP1
;  CURRENT APPLICATION NUMBER: US/10/349,143
;  CURRENT FILING DATE:  2003-01-21
;  PRIOR APPLICATION NUMBER: US/09/422,978
;  PRIOR FILING DATE: 1999-10-20
;  PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 09/298,850
;  PRIOR FILING DATE: EARLIER FILING DATE: 1999-04-21
;  PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/109,732
;  PRIOR FILING DATE: EARLIER FILING DATE: 1998-11-23
;  PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/082,614
;  PRIOR FILING DATE: EARLIER FILING DATE: 1998-04-21
;  NUMBER OF SEQ ID NOS: 11796
; SEQ ID NO 1725
;   LENGTH: 47
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
;   FEATURE:
;   NAME/KEY: allele
;   LOCATION: 24
;   OTHER INFORMATION: 99-5951-438 : polymorphic base C or T
US-10-349-143-1725

  Query Match             33.6%;  Score 14.8;  DB 17;  Length 47;
  Best Local Similarity   59.5%;  Pred. No. 2.2e+04;
  Matches   25;  Conservative    0;  Mismatches   17;  Indels    0;  Gaps    0;

Qy          3 GGGTCCCGTTCCTTCTTAATAACCGGTCGCGGTTATTAAGAA 44
              || ||||  ||  |||| ||||     | |  ||| | | ||
Db         46 GGCTCCCCATCTCTCTTTATAARTTAGCACCATTAATTATAA 5


RESULT 24
US-09-993-346-585/c
; Sequence 585, Application US/09993346
; Publication No. US20030124530A1
;   GENERAL INFORMATION:
;        APPLICANT: Edwards, Cynthia A.
;                   Cantor, Charles R.
;                   Andrews, Beth M.
;                   Turin, Lisa M.
;                   Fry, Kirk E.
;        TITLE OF INVENTION: Sequence-Directed DNA Binding
;                            Molecules, Compositions and Methods
;        NUMBER OF SEQUENCES: 664
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Genelabs Technologies, Inc.
;             STREET: 505 Penobscot Drive
;             CITY: Redwood City
;             STATE: CA
;             COUNTRY: USA
;             ZIP: 94063
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.25
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/993,346
;             FILING DATE: 13-No. US20030124530A1-2001
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: 09/354,947
;             FILING DATE: <Unknown>
;             APPLICATION NUMBER: US 08/171,389
;             FILING DATE: 20-DEC-1993
;             APPLICATION NUMBER: US 08/123,936
;             FILING DATE: 17-SEP-1993
;             APPLICATION NUMBER: US 07/996,783
;             FILING DATE: 23-DEC-1992
;             APPLICATION NUMBER: US 07/723,618
;             FILING DATE: 27-JUN-1991
;             APPLICATION NUMBER: US 08/081,070
;             FILING DATE: 22-JUN-1993
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Brady, John F.
;             REGISTRATION NUMBER: 39,118
;             REFERENCE/DOCKET NUMBER: 4600-0175.20/G19P3D1
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (650) 324-0880
;             TELEFAX: (650) 324-0960
;   INFORMATION FOR SEQ ID NO: 585:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 50 base pairs
;             TYPE: nucleic acid
;             STRANDEDNESS: double
;             TOPOLOGY: linear
;        MOLECULE TYPE: DNA (genomic)
;        HYPOTHETICAL: NO
;        ORIGINAL SOURCE:
;             INDIVIDUAL ISOLATE: Human papilloma virus type-16 E6/E7
;                                 (start site 97)
;        SEQUENCE DESCRIPTION: SEQ ID NO: 585:
US-09-993-346-585

  Query Match             33.6%;  Score 14.8;  DB 10;  Length 50;
  Best Local Similarity   73.1%;  Pred. No. 2.2e+04;
  Matches   19;  Conservative    0;  Mismatches    7;  Indels    0;  Gaps    0;

Qy         11 TTCCTTCTTAATAACCGGTCGCGGTT 36
              | | ||  || ||||||||  |||||
Db         27 TGCTTTTATACTAACCGGTTTCGGTT 2


RESULT 25
US-10-131-827-2809
; Sequence 2809, Application US/10131827
; Publication No. US20040009479A1
; GENERAL INFORMATION:
;  APPLICANT: Wohlgemuth, Jay
;  APPLICANT:  Fry, Kirk
;  APPLICANT:  Woodward, Robert
;  APPLICANT:  Ly, Ngoc
;  TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING AND MONITORING AUTOIMMUNE
```

```
;  TITLE OF INVENTION:  CHRONIC INFLAMMATORY DISEASES
;  FILE REFERENCE: 506612000120
;  CURRENT APPLICATION NUMBER: US/10/131,827
;  CURRENT FILING DATE:  2002-09-06
;  PRIOR APPLICATION NUMBER: US 10/006,290
;  PRIOR FILING DATE: 2001-10-22
;  PRIOR APPLICATION NUMBER: US 60/296,764
;  PRIOR FILING DATE: 2001-06-08
;  NUMBER OF SEQ ID NOS: 9090
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2809
;   LENGTH: 50
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-131-827-2809

  Query Match             33.6%;  Score 14.8;  DB 17;  Length 50;
  Best Local Similarity   64.7%;  Pred. No. 2.2e+04;
  Matches   22;  Conservative    0;  Mismatches   12;  Indels    0;  Gaps    0;

Qy         11 TTCCTTCTTAATAACCGGTCGCGGTTATTAAGAA 44
              || || |   || |  | | ||| ||| ||||||
Db         10 TTTCTGCAGCATTAAAGCTGGCGCTTAATAAGAA 43


RESULT 26
US-10-098-263B-97505
; Sequence 97505, Application US/10098263B
; Publication No. US20030104410A1
; GENERAL INFORMATION:
;  APPLICANT: Mittman, Michael
;  TITLE OF INVENTION: Human Microarray
;  FILE REFERENCE: 3118.1
;  CURRENT APPLICATION NUMBER: US/10/098,263B
;  CURRENT FILING DATE:  2003-01-08
;  PRIOR APPLICATION NUMBER: 60/276,759
;  PRIOR FILING DATE: 2001-03-16
;  NUMBER OF SEQ ID NOS: 131066
;  SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
; SEQ ID NO 97505
;   LENGTH: 25
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-098-263B-97505

  Query Match             33.2%;  Score 14.6;  DB 15;  Length 25;
  Best Local Similarity   81.0%;  Pred. No. 2.3e+04;
  Matches   17;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          2 CGGGTCCCGTTCCTTCTTAAT 22
              |||||||||||| ||  || |
Db          3 CGGGTCCCGTTCGTTACTACT 23


RESULT 27
US-10-719-900-202404
; Sequence 202404, Application US/10719900
; Publication No. US20050026164A1
; GENERAL INFORMATION:
;  APPLICANT: Xue Mei Zhou
;  TITLE OF INVENTION: Methods of Genetic Analysis of Mouse
;  FILE REFERENCE: 3528.1
;  CURRENT APPLICATION NUMBER: US/10/719,900
;  CURRENT FILING DATE:  2003-11-20
;  PRIOR APPLICATION NUMBER: 60/427,808
;  PRIOR FILING DATE: 2002 11 20
;  NUMBER OF SEQ ID NOS: 982914
;  SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
; SEQ ID NO 202404
;   LENGTH: 25
;   TYPE: DNA
;   ORGANISM: Mus musculus
US-10-719-900-202404

  Query Match             33.2%;  Score 14.6;  DB 19;  Length 25;
  Best Local Similarity   81.0%;  Pred. No. 2.3e+04;
  Matches   17;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          4 GGTCCCGTTCCTTCTTAATAA 24
              |||| |  |||||||| ||||
Db          2 GGTCACTCTCCTTCTTCATAA 22


RESULT 28
US-10-719-900-728815
; Sequence 728815, Application US/10719900
; Publication No. US20050026164A1
; GENERAL INFORMATION:
;  APPLICANT: Xue Mei Zhou
;  TITLE OF INVENTION: Methods of Genetic Analysis of Mouse
;  FILE REFERENCE: 3528.1
;  CURRENT APPLICATION NUMBER: US/10/719,900
;  CURRENT FILING DATE:  2003-11-20
;  PRIOR APPLICATION NUMBER: 60/427,808
;  PRIOR FILING DATE: 2002 11 20
;  NUMBER OF SEQ ID NOS: 982914
;  SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
; SEQ ID NO 728815
;   LENGTH: 25
;   TYPE: DNA
;   ORGANISM: Mus musculus
US-10-719-900-728815

  Query Match             33.2%;  Score 14.6;  DB 19;  Length 25;
  Best Local Similarity   81.0%;  Pred. No. 2.3e+04;
  Matches   17;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy         10 GTTCCTTCTTAATAACCGGTC 30
              ||||||||| | | ||| |||
Db          1 GTTCCTTCTCACTTACCAGTC 21


RESULT 29
US-10-719-900-797256
; Sequence 797256, Application US/10719900
; Publication No. US20050026164A1
; GENERAL INFORMATION:
;  APPLICANT: Xue Mei Zhou
;  TITLE OF INVENTION: Methods of Genetic Analysis of Mouse
;  FILE REFERENCE: 3528.1
;  CURRENT APPLICATION NUMBER: US/10/719,900
;  CURRENT FILING DATE:  2003-11-20
;  PRIOR APPLICATION NUMBER: 60/427,808
;  PRIOR FILING DATE: 2002 11 20
;  NUMBER OF SEQ ID NOS: 982914
;  SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
; SEQ ID NO 797256
;   LENGTH: 25
;   TYPE: DNA
;   ORGANISM: Mus musculus
US-10-719-900-797256

  Query Match             33.2%;  Score 14.6;  DB 19;  Length 25;
  Best Local Similarity   81.0%;  Pred. No. 2.3e+04;
  Matches   17;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          4 GGTCCCGTTCCTTCTTAATAA 24
              |||||| || || ||| ||||
Db          5 GGTCCCTTTACTGCTTTATAA 25


RESULT 30
US-10-719-900-816976
; Sequence 816976, Application US/10719900
; Publication No. US20050026164A1
; GENERAL INFORMATION:
```

```
;  APPLICANT: Xue Mei Zhou
;  TITLE OF INVENTION: Methods of Genetic Analysis of Mouse
;  FILE REFERENCE: 3528.1
;  CURRENT APPLICATION NUMBER: US/10/719,900
;  CURRENT FILING DATE:  2003-11-20
;  PRIOR APPLICATION NUMBER: 60/427,808
;  PRIOR FILING DATE: 2002 11 20
;  NUMBER OF SEQ ID NOS: 982914
;  SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
; SEQ ID NO 816976
;   LENGTH: 25
;   TYPE: DNA
;   ORGANISM: Mus musculus
US-10-719-900-816976

  Query Match             33.2%;  Score 14.6;  DB 19;  Length 25;
  Best Local Similarity   81.0%;  Pred. No. 2.3e+04;
  Matches   17;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          6 TCCCGTTCCTTCTTAATAACC 26
              |||| ||||||||||  | ||
Db          1 TCCCTTTCCTTCTTATAATCC 21


RESULT 31
US-10-719-900-854921
; Sequence 854921, Application US/10719900
; Publication No. US20050026164A1
; GENERAL INFORMATION:
;  APPLICANT: Xue Mei Zhou
;  TITLE OF INVENTION: Methods of Genetic Analysis of Mouse
;  FILE REFERENCE: 3528.1
;  CURRENT APPLICATION NUMBER: US/10/719,900
;  CURRENT FILING DATE:  2003-11-20
;  PRIOR APPLICATION NUMBER: 60/427,808
;  PRIOR FILING DATE: 2002 11 20
;  NUMBER OF SEQ ID NOS: 982914
;  SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
; SEQ ID NO 854921
;   LENGTH: 25
;   TYPE: DNA
;   ORGANISM: Mus musculus
US-10-719-900-854921

  Query Match             33.2%;  Score 14.6;  DB 19;  Length 25;
  Best Local Similarity   81.0%;  Pred. No. 2.3e+04;
  Matches   17;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 GCGGGTCCCGTTCCTTCTTAA 21
              ||||||  ||  ||||  |||||
Db          5 GCGGGTGCCTTTCCACCTTAA 25


RESULT 32
US-10-719-900-896407
; Sequence 896407, Application US/10719900
; Publication No. US20050026164A1
; GENERAL INFORMATION:
;  APPLICANT: Xue Mei Zhou
;  TITLE OF INVENTION: Methods of Genetic Analysis of Mouse
;  FILE REFERENCE: 3528.1
;  CURRENT APPLICATION NUMBER: US/10/719,900
;  CURRENT FILING DATE:  2003-11-20
;  PRIOR APPLICATION NUMBER: 60/427,808
;  PRIOR FILING DATE: 2002 11 20
;  NUMBER OF SEQ ID NOS: 982914
;  SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
; SEQ ID NO 896407
;   LENGTH: 25
;   TYPE: DNA
;   ORGANISM: Mus musculus
US-10-719-900-896407

  Query Match             33.2%;  Score 14.6;  DB 19;  Length 25;
  Best Local Similarity   81.0%;  Pred. No. 2.3e+04;
  Matches   17;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy         10 GTTCCTTCTTAATAACCGGTC 30
              |||||| | ||||| | ||||
Db          3 GTTCCTACGTAATATCGGGTC 23


RESULT 33
US-10-489-739-7/c
; Sequence 7, Application US/10489739
; Publication No. US20050070690A1
; GENERAL INFORMATION:
;  APPLICANT: The University of Bristol
;  APPLICANT:  Dawbarn, David
;  APPLICANT:  Allen, Shelley Jane
;  APPLICANT:  Robertson, Alan George Simpson
;  TITLE OF INVENTION: Polypeptide Purification Method
;  FILE REFERENCE: 62637.000006
;  CURRENT APPLICATION NUMBER: US/10/489,739
;  CURRENT FILING DATE:  2004-03-16
;  PRIOR APPLICATION NUMBER: GB0122400.5
;  PRIOR FILING DATE: 2001-09-17
;  PRIOR APPLICATION NUMBER: PCT/GB02/04214
;  PRIOR FILING DATE: 2002-09-17
;  NUMBER OF SEQ ID NOS: 44
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 7
;   LENGTH: 35
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: TrkBIg2 6His Reverse Primer
US-10-489-739-7

  Query Match             33.2%;  Score 14.6;  DB 19;  Length 35;
  Best Local Similarity   73.9%;  Pred. No. 2.5e+04;
  Matches   17;  Conservative    1;  Mismatches    5;  Indels    0;  Gaps    0;

Qy         21 ATAACCGGTCGCGGTTATTAAGA 43
              | |||||||||| | : ||||| |
Db         32 AAAACCGGTCGGGYYCATTAATA 10


RESULT 34
US-10-252-155-747/c
; Sequence 747, Application US/10252155
; Publication No. US20040068096A1
; GENERAL INFORMATION:
;  APPLICANT: Bristol-Myers Squibb Company
;  TITLE OF INVENTION: HUMAN SINGLE NUCLEOTIDE POLYMORPHISMS IN ORGANIC ANION TRANSPORT
;  TITLE OF INVENTION:  MULTI-DRUG RESISTANT PROTEINS
;  FILE REFERENCE: D0152 NP
;  CURRENT APPLICATION NUMBER: US/10/252,155
;  CURRENT FILING DATE:  2002-09-20
;  PRIOR APPLICATION NUMBER: US 60/324,172
;  PRIOR FILING DATE: 2001-09-21
;  PRIOR APPLICATION NUMBER: US 60/333,700
;  PRIOR FILING DATE: 2001-11-27
;  NUMBER OF SEQ ID NOS: 783
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 747
;   LENGTH: 41
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-252-155-747

  Query Match             33.2%;  Score 14.6;  DB 17;  Length 41;
  Best Local Similarity   69.0%;  Pred. No. 2.6e+04;
  Matches   20;  Conservative    0;  Mismatches    9;  Indels    0;  Gaps    0;

Qy         10 GTTCCTTCTTAATAACCGGTCGCGGTTAT 38
```

```
              | ||||  |||  |||||  |  ||||  ||
Db         41 GCTCCTCCTTTTTAACCTCTACCGGTCAT 13


RESULT 35
US-09-991-003B-16/c
; Sequence 16, Application US/09991003B
; Patent No. US20020177125A1
; GENERAL INFORMATION:
;  APPLICANT: KAMB, Carl Alexander
;  APPLICANT:  PORITZ, Mark Aaron
;  APPLICANT:  TENG, David Heng-Fai
;  TITLE OF INVENTION: Human Rhinovirus Assays, and Compositions Therefrom
;  FILE REFERENCE: 29345/36971A
;  CURRENT APPLICATION NUMBER: US/09/991,003B
;  CURRENT FILING DATE:  2002-11-16
;  NUMBER OF SEQ ID NOS: 28
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 16
;   LENGTH: 45
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: rh1D.R2 primer
US-09-991-003B-16

  Query Match             33.2%;  Score 14.6;  DB 9;  Length 45;
  Best Local Similarity   62.2%;  Pred. No. 2.6e+04;
  Matches   23;  Conservative    0;  Mismatches   14;  Indels    0;  Gaps    0;

Qy          4 GGTCCCGTTCCTTCTTAATAACCGGTCGCGGTTATTA 40
              |||  | ||   || || ||| ||| ||| ||  | |
Db         37 GGTGACATTAAATCCTATTAAGCGGCCGCTGTATTGA 1


RESULT 36
US-10-349-143-3415
; Sequence 3415, Application US/10349143
; Publication No. US20040005584A1
; GENERAL INFORMATION:
;  APPLICANT: Cohen, Daniel
;  APPLICANT:  Blumenfeld, Marta
;  APPLICANT:  Chumakov, Ilya
;  TITLE OF INVENTION: Biallelic markers for use in constructing a high density...
;  FILE REFERENCE: GENSET.020CP1
;  CURRENT APPLICATION NUMBER: US/10/349,143
;  CURRENT FILING DATE:  2003-01-21
;  PRIOR APPLICATION NUMBER: US/09/422,978
;  PRIOR FILING DATE: 1999-10-20
;  PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 09/298,850
;  PRIOR FILING DATE: EARLIER FILING DATE: 1999-04-21
;  PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/109,732
;  PRIOR FILING DATE: EARLIER FILING DATE: 1998-11-23
;  PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/082,614
;  PRIOR FILING DATE: EARLIER FILING DATE: 1998-04-21
;  NUMBER OF SEQ ID NOS: 11796
; SEQ ID NO 3415
;   LENGTH: 47
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
;   FEATURE:
;   NAME/KEY: allele
;   LOCATION: 24
;   OTHER INFORMATION: 99-3812-243 : polymorphic base T or G
US-10-349-143-3415

  Query Match             33.2%;  Score 14.6;  DB 17;  Length 47;
  Best Local Similarity   64.5%;  Pred. No. 2.7e+04;
  Matches   20;  Conservative    1;  Mismatches   10;  Indels    0;  Gaps    0;

Qy         10 GTTCCTTCTTAATAACCGGTCGCGGTTATTA 40
              |||| | |:||| ||||  || | | |  ||
Db         16 GTTCATCCKTAAAAACCATTCTCAGCTCCTA 46


RESULT 37
US-10-098-263B-88885/c
; Sequence 88885, Application US/10098263B
; Publication No. US20030104410A1
; GENERAL INFORMATION:
;  APPLICANT: Mittman, Michael
;  TITLE OF INVENTION: Human Microarray
;  FILE REFERENCE: 3118.1
;  CURRENT APPLICATION NUMBER: US/10/098,263B
;  CURRENT FILING DATE:  2003-01-08
;  PRIOR APPLICATION NUMBER: 60/276,759
;  PRIOR FILING DATE: 2001-03-16
;  NUMBER OF SEQ ID NOS: 131066
;  SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
; SEQ ID NO 88885
;   LENGTH: 25
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-098-263B-88885

  Query Match             32.7%;  Score 14.4;  DB 15;  Length 25;
  Best Local Similarity   75.0%;  Pred. No. 2.8e+04;
  Matches   18;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          3 GGGTCCCGTTCCTTCTTAATAACC 26
              | || ||| |||| || ||||| |
Db         25 GAGTACCGGTCCTCCTGAATAAAC 2


RESULT 38
US-10-719-900-221631
; Sequence 221631, Application US/10719900
; Publication No. US20050026164A1
; GENERAL INFORMATION:
;  APPLICANT: Xue Mei Zhou
;  TITLE OF INVENTION: Methods of Genetic Analysis of Mouse
;  FILE REFERENCE: 3528.1
;  CURRENT APPLICATION NUMBER: US/10/719,900
;  CURRENT FILING DATE:  2003-11-20
;  PRIOR APPLICATION NUMBER: 60/427,808
;  PRIOR FILING DATE: 2002 11 20
;  NUMBER OF SEQ ID NOS: 982914
;  SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
; SEQ ID NO 221631
;   LENGTH: 25
;   TYPE: DNA
;   ORGANISM: Mus musculus
US-10-719-900-221631

  Query Match             32.7%;  Score 14.4;  DB 19;  Length 25;
  Best Local Similarity   75.0%;  Pred. No. 2.8e+04;
  Matches   18;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy         10 GTTCCTTCTTAATAACCGGTCGCG 33
              |||| ||  || ||||| ||| ||
Db          2 GTTCTTTGCTACTAACCAGTCACG 25


RESULT 39
US-10-719-900-221632
; Sequence 221632, Application US/10719900
; Publication No. US20050026164A1
; GENERAL INFORMATION:
;  APPLICANT: Xue Mei Zhou
;  TITLE OF INVENTION: Methods of Genetic Analysis of Mouse
;  FILE REFERENCE: 3528.1
;  CURRENT APPLICATION NUMBER: US/10/719,900
;  CURRENT FILING DATE:  2003-11-20
;  PRIOR APPLICATION NUMBER: 60/427,808
;  PRIOR FILING DATE: 2002 11 20
;  NUMBER OF SEQ ID NOS: 982914
```

```
;  SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
; SEQ ID NO 221632
;   LENGTH: 25
;   TYPE: DNA
;   ORGANISM: Mus musculus
US-10-719-900-221632

  Query Match             32.7%;  Score 14.4;  DB 19;  Length 25;
  Best Local Similarity   75.0%;  Pred. No. 2.8e+04;
  Matches   18;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy         10 GTTCCTTCTTAATAACCGGTCGCG 33
              |||| ||  || ||||| ||| ||
Db          2 GTTCTTTGCTAGTAACCAGTCACG 25


RESULT 40
US-10-719-900-434524/c
; Sequence 434524, Application US/10719900
; Publication No. US20050026164A1
; GENERAL INFORMATION:
;  APPLICANT: Xue Mei Zhou
;  TITLE OF INVENTION: Methods of Genetic Analysis of Mouse
;  FILE REFERENCE: 3528.1
;  CURRENT APPLICATION NUMBER: US/10/719,900
;  CURRENT FILING DATE:  2003-11-20
;  PRIOR APPLICATION NUMBER: 60/427,808
;  PRIOR FILING DATE: 2002 11 20
;  NUMBER OF SEQ ID NOS: 982914
;  SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
; SEQ ID NO 434524
;   LENGTH: 25
;   TYPE: DNA
;   ORGANISM: Mus musculus
US-10-719-900-434524

  Query Match             32.7%;  Score 14.4;  DB 19;  Length 25;
  Best Local Similarity   75.0%;  Pred. No. 2.8e+04;
  Matches   18;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy         16 TCTTAATAACCGGTCGCGGTTATT 39
              || ||| | | ||||  |||||||
Db         25 TCCTAACAGCAGGTCTGGGTTATT 2


Search completed: May 24, 2005, 13:22:14
Job time : 340 secs
```

OM nucleic - nucleic search, using sw model

Run on: May 24, 2005, 11:54:00 ; Search time 1848 Seconds
(without alignments)
906.292 Million cell updates/sec

Title: US-10-673-063-3_COPY_900_943
Perfect score: 44
Sequence: 1 gcgggtcccgttccttctta.....ccggtcgcggttattaagaa 44

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 159776

Minimum DB seq length: 0
Maximum DB seq length: 50

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 100 summaries

Database : EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_htc:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_gss1:*
9: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	16	36.4	35	8	BZ763244	SALK_1156
	2	15.8	35.9	50	9	CR191751	CR191751 Forward s
c	3	15.4	35.0	50	9	AL754601	AL754601 Arabidops
	4	15.2	34.5	46	8	AZ438384	1M0228O24
	5	15.2	34.5	50	1	AU104874	AU104874 AU104874
	6	15.2	34.5	50	1	AU106701	AU106701 AU106701
	7	15	34.1	49	8	AZ808932	2M0072N09
	8	14.8	33.6	46	4	BJ049506	BJ049506 BJ049506
c	9	14.6	33.2	41	8	AZ795288	2M0049M22
	10	14.6	33.2	50	9	CR147596	Forward s
	11	14.4	32.7	32	8	BH811169	SALK_0575
	12	14.4	32.7	45	8	BH636447	1008011D0
	13	14.4	32.7	46	8	BH904918	SALK_1053
c	14	14.4	32.7	49	6	CB190523	pl27g03.y
	15	14.4	32.7	50	8	BH791926	SALK_0620
c	16	14.4	32.7	50	9	CR147596	Forward s
c	17	14.2	32.3	50	1	AU102908	AU102908 AU102908
	18	14	31.8	41	9	TA192A12Q	T. brucei
	19	14	31.8	48	8	AZ801189	2M0059K02
c	20	14	31.8	50	8	BH863023	SALK_0929
	21	14	31.8	50	9	CR222513	Forward s
c	22	13.8	31.4	45	8	BH636447	1008011D0
	23	13.8	31.4	46	8	BZ292212	SALK_1236
c	24	13.8	31.4	48	8	AZ831243	2M0110F20
c	25	13.8	31.4	50	1	AU104153	AU104153 AU104153
	26	13.8	31.4	50	1	AU106910	AU106910 AU106910
c	27	13.6	30.9	37	9	AL771100	AL771100 Arabidops
	28	13.6	30.9	38	1	AJ803928	AJ803928 AJ803928
	29	13.6	30.9	44	9	AG204200	AG204200 Pan trogl
	30	13.6	30.9	45	4	BJ035009	BJ035009 BJ035009
c	31	13.6	30.9	48	9	CR405537	CR405537 Arabidops
c	32	13.6	30.9	50	1	AU102726	AU102726 AU102726
	33	13.6	30.9	50	1	AU106009	AU106009 AU106009
c	34	13.4	30.5	33	9	CG724001	CG724001 1119079B0
	35	13.4	30.5	35	9	CL436566	CL436566 PST3252-N
	36	13.4	30.5	38	8	AZ642621	AZ642621 1M0505B13
c	37	13.4	30.5	39	1	AU010606	AU010606 AU010606
c	38	13.4	30.5	39	1	AU011134	AU011134 AU011134
c	39	13.4	30.5	39	1	AU012105	AU012105 AU012105
c	40	13.4	30.5	39	1	AU012381	AU012381 AU012381
c	41	13.4	30.5	39	1	AU012382	AU012382 AU012382
	42	13.4	30.5	40	8	BH796426	BH796426 1008094B1
	43	13.4	30.5	45	9	TA3G04P	AL451566 T. brucei
	44	13.4	30.5	48	7	D18810	D18810 MUSGS00961
	45	13.4	30.5	50	1	AU107655	AU107655 AU107655
	46	13.4	30.5	50	9	CR236244	CR236244 Forward s
c	47	13.2	30.0	29	1	AU256798	AU256798 AU256798
c	48	13.2	30.0	38	8	BH910864	BH910864 SALK_0629
	49	13.2	30.0	38	9	TA360D04P	AL494047 T. brucei
	50	13.2	30.0	40	8	AZ310280	AZ310280 1M0025K10
c	51	13.2	30.0	42	8	AZ665709	AZ665709 1M0547K13
c	52	13.2	30.0	43	1	AI765730	AI765730 wi83g12.x
c	53	13.2	30.0	43	1	AA196741	AA196741 zq09d02.s
	54	13.2	30.0	44	9	CL528318	CL528318 ASV14E08.
c	55	13.2	30.0	46	1	AI669689	AI669689 wc12d11.x
c	56	13.2	30.0	46	4	BJ049506	BJ049506 BJ049506
c	57	13.2	30.0	46	7	W58703	W58703 zd23a04.r1
	58	13.2	30.0	47	8	AZ783792	AZ783792 2M0025N23
c	59	13.2	30.0	47	8	BH848002	BH848002 SALK_0673
c	60	13.2	30.0	48	1	AU266867	AU266867 AU266867
c	61	13.2	30.0	48	9	AG204400	AG204400 Pan trogl
c	62	13.2	30.0	49	1	AU265180	AU265180 AU265180
	63	13.2	30.0	49	8	AQ034111	AQ034111 l(2)k0752
c	64	13.2	30.0	50	1	AU102858	AU102858 AU102858
c	65	13.2	30.0	50	1	AU104031	AU104031 AU104031
c	66	13.2	30.0	50	1	AU104033	AU104033 AU104033
c	67	13.2	30.0	50	1	AU104035	AU104035 AU104035
c	68	13.2	30.0	50	1	AU104036	AU104036 AU104036
c	69	13.2	30.0	50	1	AU104037	AU104037 AU104037
c	70	13.2	30.0	50	1	AU104038	AU104038 AU104038
c	71	13.2	30.0	50	1	AU104039	AU104039 AU104039
c	72	13.2	30.0	50	1	AU104254	AU104254 AU104254
	73	13	29.5	28	6	CA795703	CA795703 Cac_BL_27
c	74	13	29.5	31	1	AJ789302	AJ789302 AJ789302
c	75	13	29.5	32	7	D21043	D21043 HUMGS02027
	76	13	29.5	35	4	BJ047401	BJ047401 BJ047401
c	77	13	29.5	37	1	AJ666664	AJ666664 AJ666664
c	78	13	29.5	37	8	AZ592382	AZ592382 1M0403G13
	79	13	29.5	42	9	AL757728	AL757728 Arabidops
	80	13	29.5	44	7	T17569	T17569 mps v3 The
c	81	13	29.5	45	1	AA676774	AA676774 zj71f12.s
	82	13	29.5	48	6	CA968570	CA968570 CcLL02a07
	83	13	29.5	49	1	AL643698	AL643698 AL643698
	84	13	29.5	49	9	CG719534	CG719534 1119058A1
c	85	13	29.5	50	1	AU010193	AU010193 AU010193
	86	13	29.5	50	1	AU105429	AU105429 AU105429
	87	13	29.5	50	1	AU105439	AU105439 AU105439
	88	13	29.5	50	1	AU105441	AU105441 AU105441
	89	13	29.5	50	1	AU105446	AU105446 AU105446
	90	13	29.5	50	1	AU107549	AU107549 AU107549
	91	12.8	29.1	23	8	BH811030	BH811030 SALK_0571
c	92	12.8	29.1	23	9	AJ587643	AJ587643 Arabidops
c	93	12.8	29.1	28	4	BM395440	BM395440 50072-2-9
c	94	12.8	29.1	32	1	AU014466	AU014466 AU014466
	95	12.8	29.1	36	8	AZ604700	AZ604700 1M0425J10
	96	12.8	29.1	38	8	BZ660720	BZ660720 SALK_0241
	97	12.8	29.1	38	8	BZ660721	BZ660721 SALK_0241

c 98	12.8	29.1	39	2	AV965342	AV965342 AV965342
c 99	12.8	29.1	42	8	AZ492152	AZ492152 1M0326B09
c 100	12.8	29.1	43	1	AI801542	AI801542 to90g08.x

ALIGNMENTS

RESULT 1
BZ763244
LOCUS BZ763244 35 bp DNA linear GSS 13-MAR-2003
DEFINITION SALK_115680.48.40.x Arabidopsis thaliana TDNA insertion lines
Arabidopsis thaliana genomic clone SALK_115680.48.40.x, genomic survey sequence.
ACCESSION BZ763244
VERSION BZ763244.1 GI:28935797
KEYWORDS GSS.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE 1 (bases 1 to 35)
AUTHORS Alonso,J.M., Leisse,T.J., Barajas,P., Chen,H., Cheuk,R., Gadrinab,C., Jeske,A., Karnes,M., Kim,C.J., Parker,H., Prednis,L., Shinn,P., Zimmerman,J. and Ecker,J.R.
TITLE A Sequence-Indexed Library of Insertion Mutations in the Arabidopsis Genome
JOURNAL Unpublished (2001)
COMMENT Contact: Joseph R. Ecker
Salk Institute Genomic Analysis Laboratory (SIGnAL)
The Salk Institute for Biological Studies
10010 N. Torrey Pines Road, La Jolla, CA 92037, USA
Tel: 858 453 4100 x1752
Fax: 858 558 6379
Email: ecker@salk.edu
This is single pass sequence recovered from the left border of TDNA.
Class: TDNA tagged.
FEATURES Location/Qualifiers
source 1. .35
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/ecotype="Col-0"
/db_xref="taxon:3702"
/clone="SALK_115680.48.40.x"
/clone_lib="Arabidopsis thaliana TDNA insertion lines"
/note="PCR was performed on Arabidopsis thaliana lines each of which contains one or more TDNA insertion elements. The resultant fragment for each line was directly sequenced to determine the genomic sequence at the site of insertion. Details of the protocols used can be found at http://signal.salk.edu/tdna_protocols.html"
ORIGIN

Query Match 36.4%; Score 16; DB 8; Length 35;
Best Local Similarity 68.8%; Pred. No. 5.2e+04;
Matches 22; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

```
Qy          9 CGTTCCTTCTTAATAACCGGTCGCGGTTATTA 40
              ||||  |   || |||  ||||| |||| |||
Db          1 CGTTTTTCTCTACTAATTGGTCGAGGTTGTTA 32
```

RESULT 2
CR191751
LOCUS CR191751 50 bp DNA linear GSS 06-JUL-2004
DEFINITION Forward strand read from insert in 5'HPRT insertion targeting and chromosome engineering clone MHPN57o19, genomic survey sequence.
ACCESSION CR191751
VERSION CR191751.1 GI:49970600
KEYWORDS GSS; genome survey sequence; MICER.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 50)
AUTHORS Adams,D.J., Biggs,P.J., Cox,A.V., Davies,R.M., van der Weyden,L., Jonkers,J., Smith,J., Plumb,R.W., Taylor,R.G., Nishijima,I., Yu,Y., Rogers,J. and Bradley,A.
TITLE Direct Submission
JOURNAL Submitted (20-FEB-2004) Sanger Centre, Hinxton, Cambridgeshire, CB10 1SA, UK. http://www.sanger.ac.uk/MICER
FEATURES Location/Qualifiers
source 1. .50
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/clone="MHPN57o19"
/clone_lib="MHPN"
ORIGIN

Query Match 35.9%; Score 15.8; DB 9; Length 50;
Best Local Similarity 65.7%; Pred. No. 6.6e+04;
Matches 23; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

```
Qy          2 CGGGTCCCGTTCCTTCTTAATAACCGGTCGCGGTT 36
              |||| ||| |||| |   || ||| ||  | || |
Db         16 CGGGCCCCCTTCCCTTACAAAAACGGGGAGGGGCT 50
```

RESULT 3
AL754601/c
LOCUS AL754601 50 bp DNA linear GSS 01-APR-2004
DEFINITION Arabidopsis thaliana T-DNA flanking sequence GK-054H09-012427, genomic survey sequence.
ACCESSION AL754601
VERSION AL754601.1 GI:21487099
KEYWORDS GSS.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE 1
AUTHORS Li,Y., Rosso,M.G., Strizhov,N., Viehoever,P. and Weisshaar,B.
TITLE GABI-Kat SimpleSearch: a flanking sequence tag (FST) database for the identification of T-DNA insertion mutants in Arabidopsis thaliana
JOURNAL Bioinformatics 19 (11), 1441-1442 (2003)
MEDLINE 22755829
PUBMED 12874060
REFERENCE 2
AUTHORS Rosso,M.G., Li,Y., Strizhov,N., Reiss,B., Dekker,K. and Weisshaar,B.
TITLE An Arabidopsis thaliana T-DNA mutagenized population (GABI-Kat) for flanking sequence tag-based reverse genetics
JOURNAL Plant Mol. Biol. 53 (1-2), 247-259 (2003)
MEDLINE 23117147
PUBMED 14756321
REFERENCE 3
AUTHORS Strizhov,N., Li,Y., Rosso,M.G., Viehoever,P., Dekker,K.A. and Weisshaar,B.
TITLE High-throughput generation of sequence indexes from T-DNA mutagenized Arabidopsis thaliana lines
JOURNAL BioTechniques 35 (6), 1164-1168 (2003)
PUBMED 14682050
REFERENCE 4 (bases 1 to 50)
AUTHORS Rosso,M.G., Li,Y., Strizhov,N. and Weisshaar,B.
TITLE Direct Submission
JOURNAL Submitted (31-MAR-2004) Weisshaar B., Max-Planck-Institut fuer Zuechtungsforschung, Carl-von-Linne-Weg 10, Koeln, 50829, Germany
COMMENT This sequence has been recovered from the left border of the T-DNA. It indicates an insertion close to or within gene At3g62160. Details on the protocols used for generation of the sequence are described in References 1-3. The sequences are generated at the MPI

for Plant Breeding Research in the context of the GABI-Kat project. GABI-Kat is part of the German Plant Genomics program designated 'GABI'. Information on line availability can be found at: http://www.mpiz-koeln.mpg.de/GABI-Kat/.

```
FEATURES             Location/Qualifiers
     source          1. .50
                     /organism="Arabidopsis thaliana"
                     /mol_type="genomic DNA"
                     /strain="Columbia 0"
                     /db_xref="taxon:3702"
                     /clone="GK-054H09-012427"
                     /clone_lib="Arabidopsis thaliana T-DNA insertion lines"
                     /ecotype="Col-0"
                     /note="PCR was performed on DNA from Arabidopsis thaliana
                     plants (T1) which were transformed with the T-DNA from
                     vector pAC161 (GenBank accession number: AJ537514). The
                     lines contain one or more T-DNA insertions. The DNA
                     fragment(s) resulting from the PCR were directly sequenced
                     to determine the genomic sequence flanking the insertion.
                     T-DNA derived sequences were removed."
ORIGIN

  Query Match             35.0%;  Score 15.4;  DB 9;  Length 50;
  Best Local Similarity   76.0%;  Pred. No. 9.8e+04;
  Matches   19;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy         11 TTCCTTCTTAATAACCGGTCGCGGT 35
              ||||||||||  || || |||  ||
Db         29 TTCCTTCTTATCAATCGCTCGATGT 5


RESULT 4
AZ438384
LOCUS       AZ438384                 46 bp    DNA     linear   GSS 03-OCT-2000
DEFINITION  1M0228O24F Mouse 10kb plasmid UUGC1M library Mus musculus genomic
            clone UUGC1M0228O24 F, genomic survey sequence.
ACCESSION   AZ438384
VERSION     AZ438384.1  GI:10562397
KEYWORDS    GSS.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 46)
  AUTHORS   Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
            Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T.,
            Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von
            Niederhausern,A. and Wright,D.,Weiss,R.
  TITLE     Mouse whole genome scaffolding with paired end reads from 10kb
            plasmid inserts
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Robert B. Weiss
            University of Utah Genome Center
            University of Utah
            Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
            84112, USA
            Tel: 801 585 5606
            Fax: 801 585 7177
            Email: ddunn@genetics.utah.edu
            Insert Length: 10000   Std Error: 0.00
            Plate: 0228  row: O  column: 24
            Seq primer: CGTTGTAAAACGACGGCCAGT
            Class: plasmid ends
            High quality sequence stop: 46.
FEATURES             Location/Qualifiers
     source          1. .46
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="UUGC1M0228O24"
                     /sex="Male"
                     /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
                     /clone_lib="Mouse 10kb plasmid UUGC1M library"
                     /note="Vector: PWD42nv; Purified genomic DNA from M.
                     musculus C57BL/6J (male) was obtained  from the Jackson
                     Laboratory Mouse DNA Resource
                     (http://www.jax.org/resources/documents/dnares/). The DNA
                     was hydrodynamically sheared by repeated passage through a
                     0.005 inch orifice at constant velocity. The sheared DNA
                     was blunt end-repaired with T4 DNA polymerase and T4
                     polynucleotide kinase. Adaptor oligonucleotides were
                     ligated to the blunt ends in high  molar excess. The
                     adaptored DNA was purified and size-selected for a 9.5 to
                     10.5 kb range using preparative agarose gel
                     electrophoresis. Vector DNA was prepared from a derivative
                     of pWD42 (gi|4732114|gb|AF129072.1), a copy-number
                     inducible derivative of plasmid R1. The vector was ligated
                     with adaptors complementary to the insert adaptors and
                     purified. The sheared, adaptored mouse DNA was annealed to
                     adaptored vector DNA, and transformed into
                     chemically-competent E. coli XL10-Gold (Stratagene) cells
                     and selected for ampicillin resistance."
ORIGIN

  Query Match             34.5%;  Score 15.2;  DB 8;  Length 46;
  Best Local Similarity   85.0%;  Pred. No. 1.2e+05;
  Matches   17;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          7 CCCGTTCCTTCTTAATAACC 26
              ||  ||||||||| ||||||
Db          3 CCTTTTCCTTCTTTATAACC 22


RESULT 5
AU104874
LOCUS       AU104874                 50 bp    mRNA    linear   EST 28-JAN-2004
DEFINITION  AU104874 Sugano Homo sapiens cDNA library Homo sapiens cDNA clone
            HEP03014, mRNA sequence.
ACCESSION   AU104874
VERSION     AU104874.1  GI:13554395
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 50)
  AUTHORS   Suzuki,Y., Taira,H., Tsunoda,T., Mizushima-Sugano,J., Sese,J.,
            Hata,H., Ota,T., Isogai,T., Tanaka,T., Morishita,S., Okubo,K.,
            Sakaki,Y., Nakamura,Y., Suyama,A. and Sugano,S.
  TITLE     Diverse transcriptional initiation revealed by fine, large-scale
            mapping of mRNA start sites
  JOURNAL   EMBO Rep. 2 (5), 388-393 (2001)
  MEDLINE   21270072
   PUBMED   11375929
COMMENT     Contact: Yutaka Suzuki
            Department of Virology
            Institute of Medical Science, University of Tokyo
            4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
            Email: ysuzuki@ims.u-tokyo.ac.jp
            Suzuki,Y., Yoshitomo-Nakagawa,K., Maruyama,K., Suyama,A. and
            Sugano,S. Construction and characterization of a full
            length-enriched and a 5'-end-enriched cDNA library. Gene 200 (1-2),
            149-156 (1997).
FEATURES             Location/Qualifiers
     source          1. .50
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="HEP03014"
                     /clone_lib="Sugano Homo sapiens cDNA library"
ORIGIN

  Query Match             34.5%;  Score 15.2;  DB 1;  Length 50;
  Best Local Similarity   63.9%;  Pred. No. 1.2e+05;
  Matches   23;  Conservative    0;  Mismatches   13;  Indels    0;  Gaps    0;
```

```
Qy          5 GTCCCGTTCCTTCTTAATAACCGGTCGCGGTTATTA 40
              ||||| ||||| ||   |||| ||  | || ||
Db         13 GTCCCTGTCCTTTGTACACACCGCCCGTCGCTACTA 48
```

```
RESULT 6
AU106701
LOCUS       AU106701                 50 bp    mRNA    linear   EST 28-JAN-2004
DEFINITION  AU106701 Sugano Homo sapiens cDNA library Homo sapiens cDNA clone
            KAT09190, mRNA sequence.
ACCESSION   AU106701
VERSION     AU106701.1  GI:13556222
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 50)
  AUTHORS   Suzuki,Y., Taira,H., Tsunoda,T., Mizushima-Sugano,J., Sese,J.,
            Hata,H., Ota,T., Isogai,T., Tanaka,T., Morishita,S., Okubo,K.,
            Sakaki,Y., Nakamura,Y., Suyama,A. and Sugano,S.
  TITLE     Diverse transcriptional initiation revealed by fine, large-scale
            mapping of mRNA start sites
  JOURNAL   EMBO Rep. 2 (5), 388-393 (2001)
  MEDLINE   21270072
   PUBMED   11375929
COMMENT     Contact: Yutaka Suzuki
            Department of Virology
            Institute of Medical Science, University of Tokyo
            4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
            Email: ysuzuki@ims.u-tokyo.ac.jp
            Suzuki,Y., Yoshitomo-Nakagawa,K., Maruyama,K., Suyama,A. and
            Sugano,S. Construction and characterization of a full
            length-enriched and a 5'-end-enriched cDNA library. Gene 200 (1-2),
            149-156 (1997).
FEATURES             Location/Qualifiers
     source          1. .50
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="KAT09190"
                     /clone_lib="Sugano Homo sapiens cDNA library"
ORIGIN
```

```
  Query Match             34.5%;  Score 15.2;  DB 1;  Length 50;
  Best Local Similarity   71.4%;  Pred. No. 1.2e+05;
  Matches   20;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Qy         10 GTTCCTTCTTAATAACCGGTCGCGGTTA 37
              ||  ||| |||| || | | ||| ||||
Db         13 GTCGCTTTTTAACAAGCTGACGCTGTTA 40
```

```
RESULT 7
AZ808932
LOCUS       AZ808932                 49 bp    DNA     linear   GSS 20-FEB-2001
DEFINITION  2M0072N09R Mouse 10kb plasmid UUGC1M library Mus musculus genomic
            clone UUGC2M0072N09 R, genomic survey sequence.
ACCESSION   AZ808932
VERSION     AZ808932.1  GI:12974784
KEYWORDS    GSS.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 49)
  AUTHORS   Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
            Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T.,
            Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von
            Niederhausern,A. and Wright,D.,Weiss,R.
  TITLE     Mouse whole genome scaffolding with paired end reads from 10kb
            plasmid inserts
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Robert B. Weiss
            University of Utah Genome Center
            University of Utah
            Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
            84112, USA
            Tel: 801 585 5606
            Fax: 801 585 7177
            Email: ddunn@genetics.utah.edu
            Insert Length: 10000   Std Error: 0.00
            Plate: 0072  row: N  column: 09
            Seq primer: CACACAGGAAACAGCTATGACC
            Class: plasmid ends
            High quality sequence stop: 49.
FEATURES             Location/Qualifiers
     source          1. .49
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="UUGC2M0072N09"
                     /sex="Male"
                     /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
                     /clone_lib="Mouse 10kb plasmid UUGC1M library"
                     /note="Vector: PWD42nv; Purified genomic DNA from M.
                     musculus C57BL/6J (male) was obtained  from the Jackson
                     Laboratory Mouse DNA Resource
                     (http://www.jax.org/resources/documents/dnares/). The DNA
                     was hydrodynamically sheared by repeated passage through a
                     0.005 inch orifice at constant velocity. The sheared DNA
                     was blunt end-repaired with T4 DNA polymerase and T4
                     polynucleotide kinase. Adaptor oligonucleotides were
                     ligated to the blunt ends in high  molar excess. The
                     adaptored DNA was purified and size-selected for a 9.5 to
                     10.5 kb range using preparative agarose gel
                     electrophoresis. Vector DNA was prepared from a derivative
                     of pWD42 (gi|4732114|gb|AF129072.1), a copy-number
                     inducible derivative of plasmid R1. The vector was ligated
                     with adaptors complementary to the insert adaptors and
                     purified. The sheared, adaptored mouse DNA was annealed to
                     adaptored vector DNA, and transformed into
                     chemically-competent E. coli XL10-Gold (Stratagene) cells
                     and selected for ampicillin resistance."
ORIGIN
```

```
  Query Match             34.1%;  Score 15;  DB 8;  Length 49;
  Best Local Similarity   67.7%;  Pred. No. 1.4e+05;
  Matches   21;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy          7 CCCGTTCCTTCTTAATAACCGGTCGCGGTTA 37
              ||||| |||||||||| |  |  | |  | ||
Db         18 CCCGTACCTTCTTAAAAGTCTATAGGAGCTA 48
```

```
RESULT 8
BJ049506
LOCUS       BJ049506                 46 bp    mRNA    linear   EST 29-SEP-2003
DEFINITION  BJ049506 NIBB Mochii normalized Xenopus neurula library Xenopus
            laevis cDNA clone XL026f15 3', mRNA sequence.
ACCESSION   BJ049506
VERSION     BJ049506.1  GI:17378905
KEYWORDS    EST.
SOURCE      Xenopus laevis (African clawed frog)
  ORGANISM  Xenopus laevis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
            Xenopodinae; Xenopus; Xenopus.
REFERENCE   1  (bases 1 to 46)
  AUTHORS   Kitayama,A., Terasaka,C., Mochii,M., Ueno,N., Shin-i,T. and
            Kohara,Y.
  TITLE     Expressed genes in X. laevis embryo
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Tadasu Shin-i
```

```
           Center For Genetic Resource Information
           National Institute of Genetics
           1111 Yata, Mishima, Shizuoka 411-8540, Japan
           Tel: 81-559-81-6856
           Fax: 81-559-81-6855
           Email: tshini@genes.nig.ac.jp
           The information of this clone is available through the following
           URL.
           http://xenopus.nibb.ac.jp.
FEATURES             Location/Qualifiers
     source          1. .46
                     /organism="Xenopus laevis"
                     /mol_type="mRNA"
                     /db_xref="taxon:8355"
                     /clone="XL026f15"
                     /tissue_type="whole embryo"
                     /dev_stage="stage 15"
                     /clone_lib="NIBB Mochii normalized Xenopus neurula
                     library"
ORIGIN

  Query Match             33.6%;  Score 14.8;  DB 4;  Length 46;
  Best Local Similarity   67.9%;  Pred. No. 1.7e+05;
  Matches   19;  Conservative    0;  Mismatches    9;  Indels    0;  Gaps    0;

Qy         17 CTTAATAACCGGTCGCGGTTATTAAGAA 44
              | ||||||| |  |  | ||| || |||
Db         13 CATAATAACAGTACTTGNTTANTATGAA 40


RESULT 9
AZ795288/c
LOCUS       AZ795288                 41 bp    DNA     linear   GSS 16-FEB-2001
DEFINITION  2M0049M22F Mouse 10kb plasmid UUGC1M library Mus musculus genomic
            clone UUGC2M0049M22 F, genomic survey sequence.
ACCESSION   AZ795288
VERSION     AZ795288.1  GI:12942154
KEYWORDS    GSS.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 41)
  AUTHORS   Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
            Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T.,
            Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von
            Niederhausern,A. and Wright,D.,Weiss,R.
  TITLE     Mouse whole genome scaffolding with paired end reads from 10kb
            plasmid inserts
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Robert B. Weiss
            University of Utah Genome Center
            University of Utah
            Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
            84112, USA
            Tel: 801 585 5606
            Fax: 801 585 7177
            Email: ddunn@genetics.utah.edu
            Insert Length: 10000   Std Error: 0.00
            Plate: 0049  row: M  column: 22
            Seq primer: CGTTGTAAAACGACGGCCAGT
            Class: plasmid ends
            High quality sequence stop: 41.
FEATURES             Location/Qualifiers
     source          1. .41
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="UUGC2M0049M22"
                     /sex="Male"
                     /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
                     /clone_lib="Mouse 10kb plasmid UUGC1M library"
                     /note="Vector: PWD42nv; Purified genomic DNA from M.
                     musculus C57BL/6J (male) was obtained  from the Jackson
                     Laboratory Mouse DNA Resource
                     (http://www.jax.org/resources/documents/dnares/). The DNA
                     was hydrodynamically sheared by repeated passage through a
                     0.005 inch orifice at constant velocity. The sheared DNA
                     was blunt end-repaired with T4 DNA polymerase and T4
                     polynucleotide kinase. Adaptor oligonucleotides were
                     ligated to the blunt ends in high  molar excess. The
                     adaptored DNA was purified and size-selected for a 9.5 to
                     10.5 kb range using preparative agarose gel
                     electrophoresis. Vector DNA was prepared from a derivative
                     of pWD42 (gi|4732114|gb|AF129072.1), a copy-number
                     inducible derivative of plasmid R1. The vector was ligated
                     with adaptors complementary to the insert adaptors and
                     purified. The sheared, adaptored mouse DNA was annealed to
                     adaptored vector DNA, and transformed into
                     chemically-competent E. coli XL10-Gold (Stratagene) cells
                     and selected for ampicillin resistance."
ORIGIN

  Query Match             33.2%;  Score 14.6;  DB 8;  Length 41;
  Best Local Similarity   69.0%;  Pred. No. 2.1e+05;
  Matches   20;  Conservative    0;  Mismatches    9;  Indels    0;  Gaps    0;

Qy          8 CCGTTCCTTCTTAATAACCGGTCGCGGTT 36
              ||  | | ||||||||||||  ||   |||
Db         38 CCTATGCATCTTAATAACCTATCTATGTT 10


RESULT 10
CR147596
LOCUS       CR147596                 50 bp    DNA     linear   GSS 06-JUL-2004
DEFINITION  Forward strand read from insert in 5'HPRT insertion targeting and
            chromosome engineering clone MHPN78d19, genomic survey sequence.
ACCESSION   CR147596
VERSION     CR147596.1  GI:49906067
KEYWORDS    GSS; genome survey sequence; MICER.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 50)
  AUTHORS   Adams,D.J., Biggs,P.J., Cox,A.V., Davies,R.M., van der Weyden,L.,
            Jonkers,J., Smith,J., Plumb,R.W., Taylor,R.G., Nishijima,I., Yu,Y.,
            Rogers,J. and Bradley,A.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-FEB-2004) Sanger Centre, Hinxton, Cambridgeshire,
            CB10 1SA, UK. http://www.sanger.ac.uk/MICER
FEATURES             Location/Qualifiers
     source          1. .50
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10090"
                     /clone="MHPN78d19"
                     /clone_lib="MHPN"
ORIGIN

  Query Match             33.2%;  Score 14.6;  DB 9;  Length 50;
  Best Local Similarity   69.0%;  Pred. No. 2.1e+05;
  Matches   20;  Conservative    0;  Mismatches    9;  Indels    0;  Gaps    0;

Qy         11 TTCCTTCTTAATAACCGGTCGCGGTTATT 39
              ||   ||||||||| |    | |||||||
Db         12 TTGTCTCTTAATAATCTTAAGAGGTTATT 40


RESULT 11
BH811169
LOCUS       BH811169                 32 bp    DNA     linear   GSS 02-MAY-2002
DEFINITION  SALK_057563 Arabidopsis thaliana TDNA insertion lines Arabidopsis
            thaliana genomic clone SALK_057563, genomic survey sequence.
ACCESSION   BH811169
```

```
VERSION     BH811169.1  GI:20389052
KEYWORDS    GSS.
SOURCE      Arabidopsis thaliana (thale cress)
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 32)
  AUTHORS   Alonso,J.M., Leisse,T.J., Barajas,P., Chen,H., Cheuk,R.,
            Gadrinab,C., Jeske,A., Karnes,M., Kim,C.J., Parker,H., Prednis,L.,
            Shinn,P., Zimmerman,J. and Ecker,J.R.
  TITLE     A Sequence-Indexed Library of Insertion Mutations in the
            Arabidopsis Genome
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Joseph R. Ecker
            Salk Institute Genomic Analysis Laboratory (SIGnAL)
            The Salk Institute for Biological Studies
            10010 N. Torrey Pines Road, La Jolla, CA 92037, USA
            Tel: 858 453 4100 x1752
            Fax: 858 558 6379
            Email: ecker@salk.edu
            This is single pass sequence recovered from the left border of
            TDNA.
            Class: TDNA tagged.
FEATURES             Location/Qualifiers
     source          1. .32
                     /organism="Arabidopsis thaliana"
                     /mol_type="genomic DNA"
                     /ecotype="Col-0"
                     /db_xref="taxon:3702"
                     /clone="SALK_057563"
                     /clone_lib="Arabidopsis thaliana TDNA insertion lines"
                     /note="PCR was performed on Arabidopsis thaliana lines
                     each of which contains  one or more TDNA insertion
                     elements. The resultant fragment for each  line was
                     directly sequenced to determine the genomic sequence at
                     the  site of insertion. Details of the protocols used can
                     be found at  http://signal.salk.edu/tdna_protocols.html"
ORIGIN

  Query Match             32.7%;  Score 14.4;  DB 8;  Length 32;
  Best Local Similarity   65.6%;  Pred. No. 2.4e+05;
  Matches   21;  Conservative    0;  Mismatches   11;  Indels    0;  Gaps    0;

Qy         11 TTCCTTCTTAATAACCGGTCGCGGTTATTAAG 42
              ||||||  || |   ||||   |||| ||| |
Db          1 TTCCTTAATATTTTGCGGTAAGGGTTGTTATG 32


RESULT 12
BH636447
LOCUS       BH636447                  45 bp    DNA     linear   GSS 14-FEB-2002
DEFINITION  1008011D08.1EL_x1 1008 - RescueMu Grid I Zea mays genomic, genomic
            survey sequence.
ACCESSION   BH636447
VERSION     BH636447.1  GI:18658684
KEYWORDS    GSS.
SOURCE      Zea mays
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 45)
  AUTHORS   Walbot,V.
  TITLE     Maize genomic sequences found using engineered RescueMu transposon
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Walbot V
            Department of Biological Sciences
            Stanford University
            855 California Ave, Palo Alto, CA 94304, USA
            Tel: 650 723 2227
            Fax: 650 725 8221
            Email: walbot@stanford.edu
            Very probable ligation site of ends cut by single endonuclease.
            Reverse complemented post-ligation sequence from source sequence.
            Plate: 1008011  row: 35
            Class: transposon-tagged.
FEATURES             Location/Qualifiers
     source          1. .45
                     /organism="Zea mays"
                     /mol_type="genomic DNA"
                     /cultivar="mixed background W23/A188/B73"
                     /db_xref="taxon:4577"
                     /tissue_type="leaf"
                     /dev_stage="adult"
                     /lab_host="DH10B"
                     /clone_lib="1008 - RescueMu Grid I"
                     /note="Organ: leaf; Vector: RescueMu (engineered from
                     pBlueScript backbone); Site_1: BamHI; Site_2: BglII;
                     RescueMu is a 4.9 kb, modified maize Mu transposon
                     designed to allow plasmid rescue from total genomic DNA.
                     Mu elements insert preferentially into transcription
                     units. For more information on RescueMu, go to the  web
                     site www.zmdb.iastate.edu and follow the links for
                     'RescueMu.' Grid I was grown at Berkeley in 2001. DNA was
                     extracted from leaf punches,  double digested using BamHI
                     and BglII, and ligated to form circular plasmids.  DH10B
                     cells were transformed and then screened on LB plates with
                     ampicillin."
ORIGIN

  Query Match             32.7%;  Score 14.4;  DB 8;  Length 45;
  Best Local Similarity   75.0%;  Pred. No. 2.5e+05;
  Matches   18;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy         21 ATAACCGGTCGCGGTTATTAAGAA 44
              |||||| || |||| | ||  |||
Db         12 ATAACCTGTAGCGGGTTTTCGGAA 35


RESULT 13
BH904918
LOCUS       BH904918                  46 bp    DNA     linear   GSS 04-SEP-2002
DEFINITION  SALK_105328.44.50.x Arabidopsis thaliana TDNA insertion lines
            Arabidopsis thaliana genomic clone SALK_105328.44.50.x, genomic
            survey sequence.
ACCESSION   BH904918
VERSION     BH904918.1  GI:22717592
KEYWORDS    GSS.
SOURCE      Arabidopsis thaliana (thale cress)
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 46)
  AUTHORS   Alonso,J.M., Leisse,T.J., Barajas,P., Chen,H., Cheuk,R.,
            Gadrinab,C., Jeske,A., Karnes,M., Kim,C.J., Parker,H., Prednis,L.,
            Shinn,P., Zimmerman,J. and Ecker,J.R.
  TITLE     A Sequence-Indexed Library of Insertion Mutations in the
            Arabidopsis Genome
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Joseph R. Ecker
            Salk Institute Genomic Analysis Laboratory (SIGnAL)
            The Salk Institute for Biological Studies
            10010 N. Torrey Pines Road, La Jolla, CA 92037, USA
            Tel: 858 453 4100 x1752
            Fax: 858 558 6379
            Email: ecker@salk.edu
            This is single pass sequence recovered from the left border of
            TDNA. This sequence lies within 300 bases of the 3' end of
            At1g25570.
            Class: TDNA tagged.
FEATURES             Location/Qualifiers
     source          1. .46
                     /organism="Arabidopsis thaliana"
                     /mol_type="genomic DNA"
```

```
                     /ecotype="Col-0"
                     /db_xref="taxon:3702"
                     /clone="SALK_105328.44.50.x"
                     /clone_lib="Arabidopsis thaliana TDNA insertion lines"
                     /note="PCR was performed on Arabidopsis thaliana lines
                     each of which contains  one or more TDNA insertion
                     elements. The resultant fragment for each  line was
                     directly sequenced to determine the genomic sequence at
                     the  site of insertion. Details of the protocols used can
                     be found at  http://signal.salk.edu/tdna_protocols.html"
ORIGIN

  Query Match             32.7%;  Score 14.4;  DB 8;  Length 46;
  Best Local Similarity   60.0%;  Pred. No. 2.5e+05;
  Matches   24;  Conservative    0;  Mismatches   16;  Indels    0;  Gaps    0;

Qy          5 GTCCCGTTCCTTCTTAATAACCGGTCGCGGTTATTAAGAA 44
              |||   |  ||||  || |   |  || ||| ||||  |||
Db          5 GTCAAATCTCTTCCTATTTCTCTTTCTCGGCTATTGCGAA 44


RESULT 14
CB190523/c
LOCUS       CB190523                 49 bp    mRNA    linear   EST 05-FEB-2003
DEFINITION  pl27g03.y1 Ancylostoma ceylanicum adult Ancylostoma ceylanicum cDNA
            5' similar to contains element MER28 repetitive element ;, mRNA
            sequence.
ACCESSION   CB190523
VERSION     CB190523.1  GI:28253915
KEYWORDS    EST.
SOURCE      Ancylostoma ceylanicum
  ORGANISM  Ancylostoma ceylanicum
            Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Strongylida;
            Ancylostomatoidea; Ancylostomatidae; Ancylostomatinae; Ancylostoma.
REFERENCE   1  (bases 1 to 49)
  AUTHORS   McCarter,J., Clifton,S., Chiapelli,B., Pape,D., Martin,J.,
            Wylie,T., Dante,M., Marra,M., Hillier,L., Kucaba,T., Theising,B.,
            Bowers,Y., Gibbons,M., Ritter,E., Bennett,J., Franklin,C.,
            Tsagareishvili,R., Ronko,I., Kennedy,S., Maguire,L., Beck,C.,
            Underwood,K., Steptoe,M., Allen,M., Person,B., Swaller,T.,
            Harvey,N., Schurk,R., Kohn,S., Shin,T., Jackson,Y., Cardenas,M.,
            McCann,R., Waterston,R. and Wilson,R.
  TITLE     The Washington Univ. Nematode EST Project, 1999
  JOURNAL   Unpublished (1999)
COMMENT     Contact: McCarter JP
            The Washington Univ. Nematode EST Project, 1999
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            Library donated by John Hawdon of The George Washington University,
            Washington DC(mtmjmh@gwumc.edu). Claire Murphy and Dr. James
            McCarter of Washington University, GSC, St. Louis, MO mass excised
            the pBluescript phagemid from the Lamda ZAP II library.
FEATURES             Location/Qualifiers
     source          1. .49
                     /organism="Ancylostoma ceylanicum"
                     /mol_type="mRNA"
                     /db_xref="taxon:53326"
                     /dev_stage="adult"
                     /lab_host="SOLR (Stratagene)"
                     /clone_lib="Ancylostoma ceylanicum adult"
                     /note="Vector: pBluescript SK + excised from Lambda ZAP II
                     (Stratagene); Site_1: XhoI; Site_2: EcoRI; Lambda ZAP II
                     Library (99% recombinants, average insert size 1500bp,
                     amplified one time (10e6 pfu) donated by John Hawdon of
                     The George Washington University, Washington
                     DC(mtmjmh@gwumc.edu). Claire Murphy and Dr. James McCarter
                     of Washington University GSC, St. Louis, MO mass excised
                     the pBluescript phagemid from the Lamda ZAP II library."
ORIGIN

  Query Match             32.7%;  Score 14.4;  DB 6;  Length 49;
  Best Local Similarity   65.6%;  Pred. No. 2.6e+05;
  Matches   21;  Conservative    0;  Mismatches   11;  Indels    0;  Gaps    0;

Qy          2 CGGGTCCCGTTCCTTCTTAATAACCGGTCGCG 33
              |||||  |||  |  ||  ||  ||  ||||  ||
Db         49 CGGGTTCCGAACTTTATTCTTACACGGTTACG 18


RESULT 15
BH791926
LOCUS       BH791926                 50 bp    DNA     linear   GSS 02-APR-2002
DEFINITION  SALK_062026.27.05.x Arabidopsis thaliana TDNA insertion lines
            Arabidopsis thaliana genomic clone SALK_062026.27.05.x, genomic
            survey sequence.
ACCESSION   BH791926
VERSION     BH791926.1  GI:19886526
KEYWORDS    GSS.
SOURCE      Arabidopsis thaliana (thale cress)
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 50)
  AUTHORS   Alonso,J.M., Leisse,T.J., Barajas,P., Chen,H., Cheuk,R.,
            Gadrinab,C., Jeske,A., Karnes,M., Kim,C.J., Parker,H., Prednis,L.,
            Shinn,P., Zimmerman,J. and Ecker,J.R.
  TITLE     A Sequence-Indexed Library of Insertion Mutations in the
            Arabidopsis Genome
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Joseph R. Ecker
            Salk Institute Genomic Analysis Laboratory (SIGnAL)
            The Salk Institute for Biological Studies
            10010 N. Torrey Pines Road, La Jolla, CA 92037, USA
            Tel: 858 453 4100 x1752
            Fax: 858 558 6379
            Email: ecker@salk.edu
            This is single pass sequence recovered from the left border of
            TDNA.
            Class: TDNA tagged.
FEATURES             Location/Qualifiers
     source          1. .50
                     /organism="Arabidopsis thaliana"
                     /mol_type="genomic DNA"
                     /ecotype="Col-0"
                     /db_xref="taxon:3702"
                     /clone="SALK_062026.27.05.x"
                     /clone_lib="Arabidopsis thaliana TDNA insertion lines"
                     /note="PCR was performed on Arabidopsis thaliana lines
                     each of which contains  one or more TDNA insertion
                     elements. The resultant fragment for each  line was
                     directly sequenced to determine the genomic sequence at
                     the  site of insertion. Details of the protocols used can
                     be found at  http://signal.salk.edu/tdna_protocols.html"
ORIGIN

  Query Match             32.7%;  Score 14.4;  DB 8;  Length 50;
  Best Local Similarity   65.6%;  Pred. No. 2.6e+05;
  Matches   21;  Conservative    0;  Mismatches   11;  Indels    0;  Gaps    0;

Qy          7 CCCGTTCCTTCTTAATAACCGGTCGCGGTTAT 38
              ||   ||||    |||||  |  ||||  ||| ||
Db         11 CCAACTCCTATGAAATAATCTGTCGTGGTGAT 42


RESULT 16
CR147596/c
LOCUS       CR147596                 50 bp    DNA     linear   GSS 06-JUL-2004
DEFINITION  Forward strand read from insert in 5'HPRT insertion targeting and
            chromosome engineering clone MHPN78d19, genomic survey sequence.
ACCESSION   CR147596
VERSION     CR147596.1  GI:49906067
KEYWORDS    GSS; genome survey sequence; MICER.
```

```
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 50)
  AUTHORS   Adams,D.J., Biggs,P.J., Cox,A.V., Davies,R.M., van der Weyden,L.,
            Jonkers,J., Smith,J., Plumb,R.W., Taylor,R.G., Nishijima,I., Yu,Y.,
            Rogers,J. and Bradley,A.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-FEB-2004) Sanger Centre, Hinxton, Cambridgeshire,
            CB10 1SA, UK. http://www.sanger.ac.uk/MICER
FEATURES             Location/Qualifiers
     source          1. .50
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10090"
                     /clone="MHPN78d19"
                     /clone_lib="MHPN"
ORIGIN

  Query Match             32.7%;  Score 14.4;  DB 9;  Length 50;
  Best Local Similarity   75.0%;  Pred. No. 2.6e+05;
  Matches   18;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy         20 AATAACCGGTCGCGGTTATTAAGA 43
              |||||||  |   | |||||||||
Db         40 AATAACCTCTTAAGATTATTAAGA 17


RESULT 17
AU102908/c
LOCUS       AU102908                  50 bp    mRNA    linear   EST 28-JAN-2004
DEFINITION  AU102908 Sugano Homo sapiens cDNA library Homo sapiens cDNA clone
            HEP13140, mRNA sequence.
ACCESSION   AU102908
VERSION     AU102908.1  GI:13552429
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 50)
  AUTHORS   Suzuki,Y., Taira,H., Tsunoda,T., Mizushima-Sugano,J., Sese,J.,
            Hata,H., Ota,T., Isogai,T., Tanaka,T., Morishita,S., Okubo,K.,
            Sakaki,Y., Nakamura,Y., Suyama,A. and Sugano,S.
  TITLE     Diverse transcriptional initiation revealed by fine, large-scale
            mapping of mRNA start sites
  JOURNAL   EMBO Rep. 2 (5), 388-393 (2001)
  MEDLINE   21270072
   PUBMED   11375929
COMMENT     Contact: Yutaka Suzuki
            Department of Virology
            Institute of Medical Science, University of Tokyo
            4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
            Email: ysuzuki@ims.u-tokyo.ac.jp
            Suzuki,Y., Yoshitomo-Nakagawa,K., Maruyama,K., Suyama,A. and
            Sugano,S. Construction and characterization of a full
            length-enriched and a 5'-end-enriched cDNA library. Gene 200 (1-2),
            149-156 (1997).
FEATURES             Location/Qualifiers
     source          1. .50
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="HEP13140"
                     /clone_lib="Sugano Homo sapiens cDNA library"
ORIGIN

  Query Match             32.3%;  Score 14.2;  DB 1;  Length 50;
  Best Local Similarity   70.4%;  Pred. No. 3.1e+05;
  Matches   19;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Qy         16 TCTTAATAACCGGTCGCGGTTATTAAG 42
              ||| |  |||| |||||  |||  |||
Db         30 TCTGACAAACCCGTCGCCATTACCAAG 4


RESULT 18
TA192A12Q
LOCUS       TA192A12Q                 41 bp    DNA     linear   GSS 13-DEC-2000
DEFINITION  T. brucei sheared genomic DNA clone 192a12, reverse sequence,
            genomic survey sequence.
ACCESSION   AL478202
VERSION     AL478202.1  GI:11842012
KEYWORDS    GSS.
SOURCE      Trypanosoma brucei
  ORGANISM  Trypanosoma brucei
            Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae;
            Trypanosoma.
REFERENCE   1  (bases 1 to 41)
  AUTHORS   Hall,N., Bowman,S., Lennard,N.J., Doggett,J., Atkin,R.,
            Chillingworth,C., Ormond,D., Harris,B., El-Sayed,N., Hou,L.,
            Melville,S.E., Rajandream,M.A. and Barrell,B.G.
  TITLE     Direct Submission
  JOURNAL   Submitted (10-DEC-2000) Trypanosoma brucei genome sequencing
            project, Sanger Centre, The Wellcome Trust Genome Campus, Hinxton,
            Cambridge CB10 1SA, E-mail: barrell@sanger.ac.uk and
            nh1@sanger.ac.uk
COMMENT     Constructed at the Institute for Genomic Research (TIGR),
            Rockville, MD. Genomic DNA isolated from a cloned population of
            Trypanosoma brucei (TREU927/4 GUTat 10.1) was mechanically sheared
            to give a tight size distribution (
            4 kb). The v + i method used for the library construction is
            described in detail in Smith, H. and Venter, J.C. (Making small
            insert libraries for whole genome shotgun sequencing projects. In
            Genome Sequencing: A Practical Approach, eds. M. Vaudin and B.
            Barrell, Oxford University Press, 1999).
            Email: nelsayed@tigr.org
            Details of T. brucei sequencing at the Sanger Centre are available
            at http://www.sanger.ac.uk/Projects/T_brucei/.
FEATURES             Location/Qualifiers
     source          1. .41
                     /organism="Trypanosoma brucei"
                     /mol_type="genomic DNA"
                     /strain="TREU927"
                     /db_xref="taxon:5691"
                     /clone="192a12"
ORIGIN

  Query Match             31.8%;  Score 14;  DB 9;  Length 41;
  Best Local Similarity   64.5%;  Pred. No. 3.7e+05;
  Matches   20;  Conservative    0;  Mismatches   11;  Indels    0;  Gaps    0;

Qy          8 CCGTTCCTTCTTAATAACCGGTCGCGGTTAT 38
              | | |||||||||  ||| | |  || | | |
Db         10 CAGGTCCTTCTTTTTAAGCTGNGGCAGGTTT 40


RESULT 19
AZ801189
LOCUS       AZ801189                  48 bp    DNA     linear   GSS 16-FEB-2001
DEFINITION  2M0059K02R Mouse 10kb plasmid UUGC1M library Mus musculus genomic
            clone UUGC2M0059K02 R, genomic survey sequence.
ACCESSION   AZ801189
VERSION     AZ801189.1  GI:12953512
KEYWORDS    GSS.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 48)
  AUTHORS   Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
            Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T.,
            Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von
            Niederhausern,A. and Wright,D.,Weiss,R.
  TITLE     Mouse whole genome scaffolding with paired end reads from 10kb
```

```
            plasmid inserts
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Robert B. Weiss
            University of Utah Genome Center
            University of Utah
            Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
            84112, USA
            Tel: 801 585 5606
            Fax: 801 585 7177
            Email: ddunn@genetics.utah.edu
            Insert Length: 10000   Std Error: 0.00
            Plate: 0059  row: K  column: 02
            Seq primer: CACACAGGAAACAGCTATGACC
            Class: plasmid ends
            High quality sequence stop: 48.
FEATURES             Location/Qualifiers
     source          1. .48
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="UUGC2M0059K02"
                     /sex="Male"
                     /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
                     /clone_lib="Mouse 10kb plasmid UUGC1M library"
                     /note="Vector: PWD42nv; Purified genomic DNA from M.
                     musculus C57BL/6J (male) was obtained  from the Jackson
                     Laboratory Mouse DNA Resource
                     (http://www.jax.org/resources/documents/dnares/). The DNA
                     was hydrodynamically sheared by repeated passage through a
                     0.005 inch orifice at constant velocity. The sheared DNA
                     was blunt end-repaired with T4 DNA polymerase and T4
                     polynucleotide kinase. Adaptor oligonucleotides were
                     ligated to the blunt ends in high  molar excess. The
                     adaptored DNA was purified and size-selected for a 9.5 to
                     10.5 kb range using preparative agarose gel
                     electrophoresis. Vector DNA was prepared from a derivative
                     of pWD42 (gi|4732114|gb|AF129072.1), a copy-number
                     inducible derivative of plasmid R1. The vector was ligated
                     with adaptors complementary to the insert adaptors and
                     purified. The sheared, adaptored mouse DNA was annealed to
                     adaptored vector DNA, and transformed into
                     chemically-competent E. coli XL10-Gold (Stratagene) cells
                     and selected for ampicillin resistance."
ORIGIN

  Query Match             31.8%;  Score 14;  DB 8;  Length 48;
  Best Local Similarity   66.7%;  Pred. No. 3.8e+05;
  Matches   20;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy          1 GCGGGTCCCGTTCCTTCTTAATAACCGGTC 30
              || ||| | ||| |||   || |||| | |
Db          7 GCAGGTTCTGTTACTTAAGAAAAACCTGCC 36


RESULT 20
BH863023/c
LOCUS       BH863023                 50 bp    DNA     linear   GSS 05-AUG-2002
DEFINITION  SALK_092995 Arabidopsis thaliana TDNA insertion lines Arabidopsis
            thaliana genomic clone SALK_092995, genomic survey sequence.
ACCESSION   BH863023
VERSION     BH863023.1  GI:22098352
KEYWORDS    GSS.
SOURCE      Arabidopsis thaliana (thale cress)
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 50)
  AUTHORS   Alonso,J.M., Leisse,T.J., Barajas,P., Chen,H., Cheuk,R.,
            Gadrinab,C., Jeske,A., Karnes,M., Kim,C.J., Parker,H., Prednis,L.,
            Shinn,P., Zimmerman,J. and Ecker,J.R.
  TITLE     A Sequence-Indexed Library of Insertion Mutations in the
            Arabidopsis Genome
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Joseph R. Ecker
            Salk Institute Genomic Analysis Laboratory (SIGnAL)
            The Salk Institute for Biological Studies
            10010 N. Torrey Pines Road, La Jolla, CA 92037, USA
            Tel: 858 453 4100 x1752
            Fax: 858 558 6379
            Email: ecker@salk.edu
            This is single pass sequence recovered from the left border of
            TDNA. This sequence lies within an annotated intron of At3g16380.
            Class: TDNA tagged.
FEATURES             Location/Qualifiers
     source          1. .50
                     /organism="Arabidopsis thaliana"
                     /mol_type="genomic DNA"
                     /ecotype="Col-0"
                     /db_xref="taxon:3702"
                     /clone="SALK_092995"
                     /clone_lib="Arabidopsis thaliana TDNA insertion lines"
                     /note="PCR was performed on Arabidopsis thaliana lines
                     each of which contains  one or more TDNA insertion
                     elements. The resultant fragment for each  line was
                     directly sequenced to determine the genomic sequence at
                     the  site of insertion. Details of the protocols used can
                     be found at  http://signal.salk.edu/tdna_protocols.html"
ORIGIN

  Query Match             31.8%;  Score 14;  DB 8;  Length 50;
  Best Local Similarity   66.7%;  Pred. No. 3.8e+05;
  Matches   20;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy         15 TTCTTAATAACCGGTCGCGGTTATTAAGAA 44
              |||||| ||||||    |   | || | |||
Db         39 TTCTTAGTAACCGAAGGTAATGATGATGAA 10


RESULT 21
CR222513
LOCUS       CR222513                 50 bp    DNA     linear   GSS 06-JUL-2004
DEFINITION  Forward strand read from insert in 5'HPRT insertion targeting and
            chromosome engineering clone MHPN255l03, genomic survey sequence.
ACCESSION   CR222513
VERSION     CR222513.1  GI:50001362
KEYWORDS    GSS; genome survey sequence; MICER.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 50)
  AUTHORS   Adams,D.J., Biggs,P.J., Cox,A.V., Davies,R.M., van der Weyden,L.,
            Jonkers,J., Smith,J., Plumb,R.W., Taylor,R.G., Nishijima,I., Yu,Y.,
            Rogers,J. and Bradley,A.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-FEB-2004) Sanger Centre, Hinxton, Cambridgeshire,
            CB10 1SA, UK. http://www.sanger.ac.uk/MICER
FEATURES             Location/Qualifiers
     source          1. .50
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10090"
                     /clone="MHPN255l03"
                     /clone_lib="MHPN"
ORIGIN

  Query Match             31.8%;  Score 14;  DB 9;  Length 50;
  Best Local Similarity   77.3%;  Pred. No. 3.8e+05;
  Matches   17;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy         11 TTCCTTCTTAATAACCGGTCGC 32
              || ||||| || ||| || |||
Db          3 TTTCTTCTAAACAACTGGCCGC 24
```

```
RESULT 22
BH636447/c
LOCUS       BH636447                 45 bp    DNA     linear   GSS 14-FEB-2002
DEFINITION  1008011D08.1EL_x1 1008 - RescueMu Grid I Zea mays genomic, genomic
            survey sequence.
ACCESSION   BH636447
VERSION     BH636447.1  GI:18658684
KEYWORDS    GSS.
SOURCE      Zea mays
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 45)
  AUTHORS   Walbot,V.
  TITLE     Maize genomic sequences found using engineered RescueMu transposon
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Walbot V
            Department of Biological Sciences
            Stanford University
            855 California Ave, Palo Alto, CA 94304, USA
            Tel: 650 723 2227
            Fax: 650 725 8221
            Email: walbot@stanford.edu
            Very probable ligation site of ends cut by single endonuclease.
            Reverse complemented post-ligation sequence from source sequence.
            Plate: 1008011  row: 35
            Class: transposon-tagged.
FEATURES             Location/Qualifiers
     source          1. .45
                     /organism="Zea mays"
                     /mol_type="genomic DNA"
                     /cultivar="mixed background W23/A188/B73"
                     /db_xref="taxon:4577"
                     /tissue_type="leaf"
                     /dev_stage="adult"
                     /lab_host="DH10B"
                     /clone_lib="1008 - RescueMu Grid I"
                     /note="Organ: leaf; Vector: RescueMu (engineered from
                     pBlueScript backbone); Site_1: BamHI; Site_2: BglII;
                     RescueMu is a 4.9 kb, modified maize Mu transposon
                     designed to allow plasmid rescue from total genomic DNA.
                     Mu elements insert preferentially into transcription
                     units. For more information on RescueMu, go to the  web
                     site www.zmdb.iastate.edu and follow the links for
                     'RescueMu.' Grid I was grown at Berkeley in 2001. DNA was
                     extracted from leaf punches,  double digested using BamHI
                     and BglII, and ligated to form circular plasmids.  DH10B
                     cells were transformed and then screened on LB plates with
                     ampicillin."
ORIGIN

  Query Match             31.4%;  Score 13.8;  DB 8;  Length 45;
  Best Local Similarity   63.6%;  Pred. No. 4.5e+05;
  Matches   21;  Conservative    0;  Mismatches   12;  Indels    0;  Gaps    0;

Qy          6 TCCCGTTCCTTCTTAATAACCGGTCGCGGTTAT 38
              |||  |  ||||  || | ||| |   ||||||
Db         44 TCCACTAGCTTCCGAAAACCCGCTACAGGTTAT 12


RESULT 23
BZ292212
LOCUS       BZ292212                 46 bp    DNA     linear   GSS 24-OCT-2002
DEFINITION  SALK_123631.34.00.x Arabidopsis thaliana TDNA insertion lines
            Arabidopsis thaliana genomic clone SALK_123631.34.00.x, genomic
            survey sequence.
ACCESSION   BZ292212
VERSION     BZ292212.1  GI:24338860
KEYWORDS    GSS.
SOURCE      Arabidopsis thaliana (thale cress)
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 46)
  AUTHORS   Alonso,J.M., Leisse,T.J., Barajas,P., Chen,H., Cheuk,R.,
            Gadrinab,C., Jeske,A., Karnes,M., Kim,C.J., Parker,H., Prednis,L.,
            Shinn,P., Zimmerman,J. and Ecker,J.R.
  TITLE     A Sequence-Indexed Library of Insertion Mutations in the
            Arabidopsis Genome
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Joseph R. Ecker
            Salk Institute Genomic Analysis Laboratory (SIGnAL)
            The Salk Institute for Biological Studies
            10010 N. Torrey Pines Road, La Jolla, CA 92037, USA
            Tel: 858 453 4100 x1752
            Fax: 858 558 6379
            Email: ecker@salk.edu
            This is single pass sequence recovered from the left border of
            TDNA.
            Class: TDNA tagged.
FEATURES             Location/Qualifiers
     source          1. .46
                     /organism="Arabidopsis thaliana"
                     /mol_type="genomic DNA"
                     /ecotype="Col-0"
                     /db_xref="taxon:3702"
                     /clone="SALK_123631.34.00.x"
                     /clone_lib="Arabidopsis thaliana TDNA insertion lines"
                     /note="PCR was performed on Arabidopsis thaliana lines
                     each of which contains  one or more TDNA insertion
                     elements. The resultant fragment for each  line was
                     directly sequenced to determine the genomic sequence at
                     the  site of insertion. Details of the protocols used can
                     be found at  http://signal.salk.edu/tdna_protocols.html"
ORIGIN

  Query Match             31.4%;  Score 13.8;  DB 8;  Length 46;
  Best Local Similarity   72.0%;  Pred. No. 4.5e+05;
  Matches   18;  Conservative    0;  Mismatches    7;  Indels    0;  Gaps    0;

Qy          4 GGTCCCGTTCCTTCTTAATAACCGG 28
              ||||||  |||||||  |||   ||
Db          5 GGTCCCAATCCTTCTCTATAGTAGG 29


RESULT 24
AZ831243/c
LOCUS       AZ831243                 48 bp    DNA     linear   GSS 20-FEB-2001
DEFINITION  2M0110F20R Mouse 10kb plasmid UUGC1M library Mus musculus genomic
            clone UUGC2M0110F20 R, genomic survey sequence.
ACCESSION   AZ831243
VERSION     AZ831243.1  GI:13001151
KEYWORDS    GSS.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 48)
  AUTHORS   Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
            Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T.,
            Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von
            Niederhausern,A. and Wright,D.,Weiss,R.
  TITLE     Mouse whole genome scaffolding with paired end reads from 10kb
            plasmid inserts
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Robert B. Weiss
            University of Utah Genome Center
            University of Utah
            Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
            84112, USA
            Tel: 801 585 5606
            Fax: 801 585 7177
            Email: ddunn@genetics.utah.edu
```

```
           Insert Length: 10000   Std Error: 0.00
           Plate: 0110  row: F  column: 20
           Seq primer: CACACAGGAAACAGCTATGACC
           Class: plasmid ends
           High quality sequence stop: 48.
FEATURES             Location/Qualifiers
     source          1. .48
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="UUGC2M0110F20"
                     /sex="Male"
                     /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
                     /clone_lib="Mouse 10kb plasmid UUGC1M library"
                     /note="Vector: PWD42nv; Purified genomic DNA from M.
                     musculus C57BL/6J (male) was obtained  from the Jackson
                     Laboratory Mouse DNA Resource
                     (http://www.jax.org/resources/documents/dnares/). The DNA
                     was hydrodynamically sheared by repeated passage through a
                     0.005 inch orifice at constant velocity. The sheared DNA
                     was blunt end-repaired with T4 DNA polymerase and T4
                     polynucleotide kinase. Adaptor oligonucleotides were
                     ligated to the blunt ends in high  molar excess. The
                     adaptored DNA was purified and size-selected for a 9.5 to
                     10.5 kb range using preparative agarose gel
                     electrophoresis. Vector DNA was prepared from a derivative
                     of pWD42 (gi|4732114|gb|AF129072.1), a copy-number
                     inducible derivative of plasmid R1. The vector was ligated
                     with adaptors complementary to the insert adaptors and
                     purified. The sheared, adaptored mouse DNA was annealed to
                     adaptored vector DNA, and transformed into
                     chemically-competent E. coli XL10-Gold (Stratagene) cells
                     and selected for ampicillin resistance."
ORIGIN

  Query Match             31.4%;  Score 13.8;  DB 8;  Length 48;
  Best Local Similarity   63.6%;  Pred. No. 4.6e+05;
  Matches   21;  Conservative    0;  Mismatches   12;  Indels    0;  Gaps    0;

Qy          8 CCGTTCCTTCTTAATAACCGGTCGCGGTTATTA 40
              || |||| ||  || | | |   ||  ||||||
Db         47 CCCTTCCATCAGAAAACCTGTCAGCATTTATTA 15


RESULT 25
AU104153/c
LOCUS       AU104153                 50 bp    mRNA    linear   EST 28-JAN-2004
DEFINITION  AU104153 Sugano Homo sapiens cDNA library Homo sapiens cDNA clone
            HEP15323, mRNA sequence.
ACCESSION   AU104153
VERSION     AU104153.1  GI:13553674
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 50)
  AUTHORS   Suzuki,Y., Taira,H., Tsunoda,T., Mizushima-Sugano,J., Sese,J.,
            Hata,H., Ota,T., Isogai,T., Tanaka,T., Morishita,S., Okubo,K.,
            Sakaki,Y., Nakamura,Y., Suyama,A. and Sugano,S.
  TITLE     Diverse transcriptional initiation revealed by fine, large-scale
            mapping of mRNA start sites
  JOURNAL   EMBO Rep. 2 (5), 388-393 (2001)
  MEDLINE   21270072
   PUBMED   11375929
COMMENT     Contact: Yutaka Suzuki
            Department of Virology
            Institute of Medical Science, University of Tokyo
            4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
            Email: ysuzuki@ims.u-tokyo.ac.jp
            Suzuki,Y., Yoshitomo-Nakagawa,K., Maruyama,K., Suyama,A. and
            Sugano,S. Construction and characterization of a full
            length-enriched and a 5'-end-enriched cDNA library. Gene 200 (1-2),
            149-156 (1997).
FEATURES             Location/Qualifiers
     source          1. .50
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="HEP15323"
                     /clone_lib="Sugano Homo sapiens cDNA library"
ORIGIN

  Query Match             31.4%;  Score 13.8;  DB 1;  Length 50;
  Best Local Similarity   72.0%;  Pred. No. 4.6e+05;
  Matches   18;  Conservative    0;  Mismatches    7;  Indels    0;  Gaps    0;

Qy          1 GCGGGTCCCGTTCCTTCTTAATAAC 25
              || ||||| |   || ||| |||||
Db         37 GCCGGTCCTGAGACTGCTTTATAAC 13


RESULT 26
AU106910
LOCUS       AU106910                 50 bp    mRNA    linear   EST 28-JAN-2004
DEFINITION  AU106910 Sugano Homo sapiens cDNA library Homo sapiens cDNA clone
            HRC01450, mRNA sequence.
ACCESSION   AU106910
VERSION     AU106910.1  GI:13556431
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 50)
  AUTHORS   Suzuki,Y., Taira,H., Tsunoda,T., Mizushima-Sugano,J., Sese,J.,
            Hata,H., Ota,T., Isogai,T., Tanaka,T., Morishita,S., Okubo,K.,
            Sakaki,Y., Nakamura,Y., Suyama,A. and Sugano,S.
  TITLE     Diverse transcriptional initiation revealed by fine, large-scale
            mapping of mRNA start sites
  JOURNAL   EMBO Rep. 2 (5), 388-393 (2001)
  MEDLINE   21270072
   PUBMED   11375929
COMMENT     Contact: Yutaka Suzuki
            Department of Virology
            Institute of Medical Science, University of Tokyo
            4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
            Email: ysuzuki@ims.u-tokyo.ac.jp
            Suzuki,Y., Yoshitomo-Nakagawa,K., Maruyama,K., Suyama,A. and
            Sugano,S. Construction and characterization of a full
            length-enriched and a 5'-end-enriched cDNA library. Gene 200 (1-2),
            149-156 (1997).
FEATURES             Location/Qualifiers
     source          1. .50
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="HRC01450"
                     /clone_lib="Sugano Homo sapiens cDNA library"
ORIGIN

  Query Match             31.4%;  Score 13.8;  DB 1;  Length 50;
  Best Local Similarity   72.0%;  Pred. No. 4.6e+05;
  Matches   18;  Conservative    0;  Mismatches    7;  Indels    0;  Gaps    0;

Qy          3 GGGTCCCGTTCCTTCTTAATAACCG 27
              ||||||| | ||| |||  || | |
Db         24 GGGTCCCCTCCCTCCTTCCTATCGG 48


RESULT 27
AL771100/c
LOCUS       AL771100                 37 bp    DNA     linear   GSS 01-APR-2004
DEFINITION  Arabidopsis thaliana T-DNA flanking sequence GK-177C09-013534,
            genomic survey sequence.
```

```
ACCESSION   AL771100
VERSION     AL771100.1  GI:21533302
KEYWORDS    GSS.
SOURCE      Arabidopsis thaliana (thale cress)
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1
  AUTHORS   Li,Y., Rosso,M.G., Strizhov,N., Viehoever,P. and Weisshaar,B.
  TITLE     GABI-Kat SimpleSearch: a flanking sequence tag (FST) database for
            the identification of T-DNA insertion mutants in Arabidopsis
            thaliana
  JOURNAL   Bioinformatics 19 (11), 1441-1442 (2003)
  MEDLINE   22755829
   PUBMED   12874060
REFERENCE   2
  AUTHORS   Rosso,M.G., Li,Y., Strizhov,N., Reiss,B., Dekker,K. and
            Weisshaar,B.
  TITLE     An Arabidopsis thaliana T-DNA mutagenized population (GABI-Kat) for
            flanking sequence tag-based reverse genetics
  JOURNAL   Plant Mol. Biol. 53 (1-2), 247-259 (2003)
  MEDLINE   23117147
   PUBMED   14756321
REFERENCE   3
  AUTHORS   Strizhov,N., Li,Y., Rosso,M.G., Viehoever,P., Dekker,K.A. and
            Weisshaar,B.
  TITLE     High-throughput generation of sequence indexes from T-DNA
            mutagenized Arabidopsis thaliana lines
  JOURNAL   BioTechniques 35 (6), 1164-1168 (2003)
   PUBMED   14682050
REFERENCE   4  (bases 1 to 37)
  AUTHORS   Li,Y., Strizhov,N., Rosso,M.G. and Weisshaar,B.
  TITLE     Direct Submission
  JOURNAL   Submitted (31-MAR-2004) Weisshaar B., Max-Planck-Institut fuer
            Zuechtungsforschung, Carl-von-Linne-Weg 10, Koeln, 50829, Germany
COMMENT     This sequence has been recovered from the left border of the T-DNA.
            It indicates an insertion within the locus defined by BAC clone
            T2K10. Details on the protocols used for generation of the sequence
            are described in References 1-3. The sequences are generated at the
            MPI for Plant Breeding Research in the context of the GABI-Kat
            project. GABI-Kat is part of the German Plant Genomics program
            designated 'GABI'. Information on line availability can be found
            at: http://www.mpiz-koeln.mpg.de/GABI-Kat/.
FEATURES             Location/Qualifiers
     source          1. .37
                     /organism="Arabidopsis thaliana"
                     /mol_type="genomic DNA"
                     /strain="Columbia 0"
                     /db_xref="taxon:3702"
                     /clone="GK-177C09-013534"
                     /clone_lib="Arabidopsis thaliana T-DNA insertion lines"
                     /ecotype="Col-0"
                     /note="PCR was performed on DNA from Arabidopsis thaliana
                     plants (T1) which were transformed with the T-DNA from
                     vector pAC161 (GenBank accession number: AJ537514). The
                     lines contain one or more T-DNA insertions. The DNA
                     fragment(s) resulting from the PCR were directly sequenced
                     to determine the genomic sequence flanking the insertion.
                     T-DNA derived sequences were removed."
ORIGIN

  Query Match             30.9%;  Score 13.6;  DB 9;  Length 37;
  Best Local Similarity   67.9%;  Pred. No. 5.3e+05;
  Matches   19;  Conservative    0;  Mismatches    9;  Indels    0;  Gaps    0;

Qy          3 GGGTCCCGTTCCTTCTTAATAACCGGTC 30
              || |||| |   |||| |||   |||||
Db         37 GGATCCCATAGGTTCTAAATTGACGGTC 10


RESULT 28
AJ803928
```

```
LOCUS       AJ803928                38 bp    mRNA    linear   EST 11-AUG-2004
DEFINITION  AJ803928 Antirrhinum majus whole plant Antirrhinum majus cDNA clone
            018_5_08_i07, mRNA sequence.
ACCESSION   AJ803928
VERSION     AJ803928.1  GI:51119256
KEYWORDS    EST.
SOURCE      Antirrhinum majus (snapdragon)
  ORGANISM  Antirrhinum majus
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            asterids; lamiids; Lamiales; Plantaginaceae; Antirrhineae;
            Antirrhinum.
REFERENCE   1  (bases 1 to 38)
  AUTHORS   Zachgo,S., Stueber,K., Saedler,H., Sommer,H. and Schwarz-Sommer,Z.
  TITLE     Antirrhinum EST collection
  JOURNAL   Unpublished (2003)
COMMENT     Contact: Schwarz-Sommer Z
            Molekulare Pflanzengenetik
            MPI fuer Zuechtungsforschung
            Carl-von-Linne Weg 10, D-50829, Germany.
FEATURES             Location/Qualifiers
     source          1. .38
                     /organism="Antirrhinum majus"
                     /mol_type="mRNA"
                     /db_xref="taxon:4151"
                     /clone="018_5_08_i07"
                     /tissue_type="whole plant"
                     /clone_lib="Antirrhinum majus whole plant"
ORIGIN

  Query Match             30.9%;  Score 13.6;  DB 1;  Length 38;
  Best Local Similarity   80.0%;  Pred. No. 5.4e+05;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          6 TCCCGTTCCTTCTTAATAAC 25
              |||| |||  || |||||||
Db          4 TCCCATTCGGTCATAATAAC 23


RESULT 29
AG204200
LOCUS       AG204200                44 bp    DNA     linear   GSS 06-MAR-2004
DEFINITION  Pan troglodytes DNA, clone: RP43-089I10.T7, genomic survey
            sequence.
ACCESSION   AG204200
VERSION     AG204200.1  GI:45236375
KEYWORDS    GSS.
SOURCE      Pan troglodytes (chimpanzee)
  ORGANISM  Pan troglodytes
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
REFERENCE   1
  AUTHORS   Park,H., Kim,Y., Kim,S., Han,Y., Woo,T., Park,K., Eun,C.J.,
            Hoon,S.T., Chu,M., Kim,H., Joo,S., Kim,C., Song,W. and Yoo,H.
  TITLE     BAC end sequences of Library RP-43
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 44)
  AUTHORS   Park,H., Kim,Y., Kim,S., Han,Y., Woo,T., Park,K., Eun,C.J.,
            Hoon,S.T., Chu,M., Kim,H., Joo,S., Kim,C., Song,W. and Yoo,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (07-JAN-2002) Hong-Seog Park, Korea Research Institute of
            Bioscience and Biotechnology (KRIBB), Genome Research Center (GRC);
            52, Oun-dong, Yusong-gu, Daejeon 305-333, Korea
            (E-mail:redstone@mail.kribb.re.kr, URL:http://phs.grc.kribb.re.kr/,
            Tel:82-42-866-7181, Fax:82-42-860-4409)
COMMENT     Clones are derived from the chimpanzee BAC library RP-43 This BAC
            end was generated during the R&D process and may have higher chance
            of clone tracking errors.
             PRIMERS
                 Sequencing: T7
             LIBRARY
                 Vector    : pBACe3.6
                 R.Site 1  : EcoRI
```

```
                R.Site 2  : EcoRI.
FEATURES             Location/Qualifiers
     source          1. .44
                     /organism="Pan troglodytes"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9598"
                     /clone="RP43-089I10.T7"
                     /sex="male"
                     /cell_type="lymphocytes"
                     /clone_lib="RP-43 Chimpanzee Male BAC Library"
ORIGIN

  Query Match             30.9%;  Score 13.6;  DB 9;  Length 44;
  Best Local Similarity   80.0%;  Pred. No. 5.5e+05;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy         25 CCGGTCGCGGTTATTAAGAA 44
              | | | | ||||||||||||
Db         12 CGGCTTGTGGTTATTAAGAA 31


RESULT 30
BJ035009
LOCUS       BJ035009                 45 bp    mRNA    linear   EST 26-SEP-2003
DEFINITION  BJ035009 NIBB Mochii normalized Xenopus neurula library Xenopus
            laevis cDNA clone XL031k05 5', mRNA sequence.
ACCESSION   BJ035009
VERSION     BJ035009.1  GI:17414262
KEYWORDS    EST.
SOURCE      Xenopus laevis (African clawed frog)
  ORGANISM  Xenopus laevis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
            Xenopodinae; Xenopus; Xenopus.
REFERENCE   1  (bases 1 to 45)
  AUTHORS   Kitayama,A., Terasaka,C., Mochii,M., Ueno,N., Shin-i,T. and
            Kohara,Y.
  TITLE     Expressed genes in X. laevis embryo
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Tadasu Shin-i
            Center For Genetic Resource Information
            National Institute of Genetics
            1111 Yata, Mishima, Shizuoka 411-8540, Japan
            Tel: 81-559-81-6856
            Fax: 81-559-81-6855
            Email: tshini@genes.nig.ac.jp
            The information of this clone is available through the following
            URL.
            http://xenopus.nibb.ac.jp.
FEATURES             Location/Qualifiers
     source          1. .45
                     /organism="Xenopus laevis"
                     /mol_type="mRNA"
                     /db_xref="taxon:8355"
                     /clone="XL031k05"
                     /tissue_type="whole embryo"
                     /dev_stage="stage 15"
                     /clone_lib="NIBB Mochii normalized Xenopus neurula
                     library"
ORIGIN

  Query Match             30.9%;  Score 13.6;  DB 4;  Length 45;
  Best Local Similarity   67.9%;  Pred. No. 5.5e+05;
  Matches   19;  Conservative    0;  Mismatches    9;  Indels    0;  Gaps    0;

Qy          7 CCCGTTCCTTCTTAATAACCGGTCGCGG 34
              | ||  |||| || | | | ||||| ||
Db         11 CTCGAGCCTTTTTTACACCAGGTCGGGG 38


RESULT 31
CR405537/c
LOCUS       CR405537                 48 bp    DNA     linear   GSS 02-MAY-2004
DEFINITION  Arabidopsis thaliana T-DNA flanking sequence GK-879C05-026471,
            genomic survey sequence.
ACCESSION   CR405537
VERSION     CR405537.1  GI:46946265
KEYWORDS    GSS.
SOURCE      Arabidopsis thaliana (thale cress)
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1
  AUTHORS   Li,Y., Rosso,M.G., Strizhov,N., Viehoever,P. and Weisshaar,B.
  TITLE     GABI-Kat SimpleSearch: a flanking sequence tag (FST) database for
            the identification of T-DNA insertion mutants in Arabidopsis
            thaliana
  JOURNAL   Bioinformatics 19 (11), 1441-1442 (2003)
  MEDLINE   22755829
   PUBMED   12874060
REFERENCE   2
  AUTHORS   Rosso,M.G., Li,Y., Strizhov,N., Reiss,B., Dekker,K. and
            Weisshaar,B.
  TITLE     An Arabidopsis thaliana T-DNA mutagenized population (GABI-Kat) for
            flanking sequence tag-based reverse genetics
  JOURNAL   Plant Mol. Biol. 53 (1-2), 247-259 (2003)
  MEDLINE   23117147
   PUBMED   14756321
REFERENCE   3
  AUTHORS   Strizhov,N., Li,Y., Rosso,M.G., Viehoever,P., Dekker,K.A. and
            Weisshaar,B.
  TITLE     High-throughput generation of sequence indexes from T-DNA
            mutagenized Arabidopsis thaliana lines
  JOURNAL   BioTechniques 35 (6), 1164-1168 (2003)
   PUBMED   14682050
REFERENCE   4  (bases 1 to 48)
  AUTHORS   Strizhov,N., Rosso,M.G., Li,Y. and Weisshaar,B.
  TITLE     Direct Submission
  JOURNAL   Submitted (01-MAY-2004) Weisshaar B., Max-Planck-Institut fuer
            Zuechtungsforschung, Carl-von-Linne-Weg 10, Koeln, 50829, Germany
COMMENT     This sequence has been recovered from the left border of the T-DNA.
            It indicates an insertion close to or within gene At2g25360.
            Details on the protocols used for generation of the sequence are
            described in References 1-3. The sequences are generated at the MPI
            for Plant Breeding Research in the context of the GABI-Kat project.
            GABI-Kat is part of the German Plant Genomics program designated
            'GABI'. Information on line availability can be found at:
            http://www.mpiz-koeln.mpg.de/GABI-Kat/.
FEATURES             Location/Qualifiers
     source          1. .48
                     /organism="Arabidopsis thaliana"
                     /mol_type="genomic DNA"
                     /strain="Columbia 0"
                     /db_xref="taxon:3702"
                     /clone="GK-879C05-026471"
                     /clone_lib="Arabidopsis thaliana T-DNA insertion lines"
                     /ecotype="Col-0"
                     /note="PCR was performed on DNA from Arabidopsis thaliana
                     plants (T1) which were transformed with the T-DNA from
                     vector pAC161 (GenBank accession number: AJ537514). The
                     lines contain one or more T-DNA insertions. The DNA
                     fragment(s) resulting from the PCR were directly sequenced
                     to determine the genomic sequence flanking the insertion.
                     T-DNA derived sequences were removed."
ORIGIN

  Query Match             30.9%;  Score 13.6;  DB 9;  Length 48;
  Best Local Similarity   65.5%;  Pred. No. 5.5e+05;
  Matches   19;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy         15 TTCTTAATAACCGGTCGCGGTTATTAAGA 43
              | | || |||| || |  ||||  || ||
Db         45 TGCCTANTAACGGGACATGGTTTATATGA 17
```

```
RESULT 32
AU102726/c
LOCUS       AU102726                 50 bp    mRNA    linear   EST 28-JAN-2004
DEFINITION  AU102726 Sugano Homo sapiens cDNA library Homo sapiens cDNA clone
            CAS08871, mRNA sequence.
ACCESSION   AU102726
VERSION     AU102726.1  GI:13552247
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 50)
  AUTHORS   Suzuki,Y., Taira,H., Tsunoda,T., Mizushima-Sugano,J., Sese,J.,
            Hata,H., Ota,T., Isogai,T., Tanaka,T., Morishita,S., Okubo,K.;
            Sakaki,Y., Nakamura,Y., Suyama,A. and Sugano,S.
  TITLE     Diverse transcriptional initiation revealed by fine, large-scale
            mapping of mRNA start sites
  JOURNAL   EMBO Rep. 2 (5), 388-393 (2001)
  MEDLINE   21270072
   PUBMED   11375929
COMMENT     Contact: Yutaka Suzuki
            Department of Virology
            Institute of Medical Science, University of Tokyo
            4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
            Email: ysuzuki@ims.u-tokyo.ac.jp
            Suzuki,Y., Yoshitomo-Nakagawa,K., Maruyama,K., Suyama,A. and
            Sugano,S. Construction and characterization of a full
            length-enriched and a 5'-end-enriched cDNA library. Gene 200 (1-2),
            149-156 (1997).
FEATURES             Location/Qualifiers
     source          1. .50
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CAS08871"
                     /clone_lib="Sugano Homo sapiens cDNA library"
ORIGIN

  Query Match             30.9%;  Score 13.6;  DB 1;  Length 50;
  Best Local Similarity   80.0%;  Pred. No. 5.6e+05;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy         25 CCGGTCGCGGTTATTAAGAA 44
              |||  |||||  ||||||||
Db         21 CCGCCCGCGGGCATTAAGAA 2


RESULT 33
AU106009
LOCUS       AU106009                 50 bp    mRNA    linear   EST 28-JAN-2004
DEFINITION  AU106009 Sugano Homo sapiens cDNA library Homo sapiens cDNA clone
            HEP22332, mRNA sequence.
ACCESSION   AU106009
VERSION     AU106009.1  GI:13555530
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 50)
  AUTHORS   Suzuki,Y., Taira,H., Tsunoda,T., Mizushima-Sugano,J., Sese,J.,
            Hata,H., Ota,T., Isogai,T., Tanaka,T., Morishita,S., Okubo,K.,
            Sakaki,Y., Nakamura,Y., Suyama,A. and Sugano,S.
  TITLE     Diverse transcriptional initiation revealed by fine, large-scale
            mapping of mRNA start sites
  JOURNAL   EMBO Rep. 2 (5), 388-393 (2001)
  MEDLINE   21270072
   PUBMED   11375929
COMMENT     Contact: Yutaka Suzuki
            Department of Virology
            Institute of Medical Science, University of Tokyo
            4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
            Email: ysuzuki@ims.u-tokyo.ac.jp
            Suzuki,Y., Yoshitomo-Nakagawa,K., Maruyama,K., Suyama,A. and
            Sugano,S. Construction and characterization of a full
            length-enriched and a 5'-end-enriched cDNA library. Gene 200 (1-2),
            149-156 (1997).
FEATURES             Location/Qualifiers
     source          1. .50
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="HEP22332"
                     /clone_lib="Sugano Homo sapiens cDNA library"
ORIGIN

  Query Match             30.9%;  Score 13.6;  DB 1;  Length 50;
  Best Local Similarity   67.9%;  Pred. No. 5.6e+05;
  Matches   19;  Conservative    0;  Mismatches    9;  Indels    0;  Gaps    0;

Qy          9 CGTTCCTTCTTAATAACCGGTCGCGGTT 36
              || | ||||  | | ||| |||| | ||
Db         15 CGCTTCTTCGAAGTTACCTGTCGGGCTT 42


RESULT 34
CG724001/c
LOCUS       CG724001                 33 bp    DNA     linear   GSS 20-OCT-2003
DEFINITION  1119079B05.2EL_x1 1119 - RescueMu Grid AA Zea mays genomic, genomic
            survey sequence.
ACCESSION   CG724001
VERSION     CG724001.1  GI:37760402
KEYWORDS    GSS.
SOURCE      Zea mays
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 33)
  AUTHORS   Walbot,V.
  TITLE     Maize genomic sequences found using engineered RescueMu transposon
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Walbot V
            Department of Biological Sciences
            Stanford University
            855 California Ave, Palo Alto, CA 94304, USA
            Tel: 650 723 2227
            Fax: 650 725 8221
            Email: walbot@stanford.edu
            Possible ligation site of ends cut by 2 different endonucleases.
            Reverse complemented post-ligation sequence from source sequence.
            Plate: 1119079  row: B  column: 05
            Class: transposon-tagged.
FEATURES             Location/Qualifiers
     source          1. .33
                     /organism="Zea mays"
                     /mol_type="genomic DNA"
                     /cultivar="mixed background W23/A188/B73/K55"
                     /db_xref="taxon:4577"
                     /tissue_type="leaf"
                     /dev_stage="adult"
                     /lab_host="DH10B"
                     /clone_lib="1119 - RescueMu Grid AA"
                     /note="Organ: leaf; Vector: RescueMu (engineered from
                     pBlueScript backbone); Site_1: BamHI; Site_2: BglII;
                     RescueMu is a 4.9 kb, modified maize Mu transposon
                     designed to allow plasmid rescue from total genomic DNA.
                     Mu elements insert preferentially into transcription
                     units. For more information on RescueMu, go to the web
                     site 'www.zmdb.iastate.edu' and follow the links for
                     'RescueMu.' Grid AA was grown at UC San Diego in 2002. DNA
                     was extracted from leaf strips, double digested using
                     BamHI and BglII, and ligated to form circular plasmids.
                     DH10B cells were transformed and then screened on LB
                     plates with ampicillin."
```

```
ORIGIN

  Query Match            30.5%;  Score 13.4;  DB 9;  Length 33;
  Best Local Similarity  73.9%;  Pred. No. 6.4e+05;
  Matches   17;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          4 GGTCCCGTTCCTTCTTAATAACC 26
              | ||||||  | ||||| || ||
Db         33 GATCCCGTCTCGTCTTAGTATCC 11


RESULT 35
CL436566
LOCUS       CL436566                35 bp    DNA     linear   GSS 18-MAR-2004
DEFINITION  PST3252-NR.Seq MICB1 Mus musculus genomic clone PST3252-NR.Seq
            similar to Gpiap1, genomic survey sequence.
ACCESSION   CL436566
VERSION     CL436566.1  GI:45571487
KEYWORDS    GSS.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 35)
  AUTHORS   Hicks,G.G.
  TITLE     www.EScells.ca
  JOURNAL   Unpublished (2002)
COMMENT     Contact: Hicks GG
            Mammalian Functional Genomics Centre
            Manitoba Institute of Cell Biology, Universitry of Manitoba
            ON5029, 675 McDermot Ave, Winnipeg, MB R3E 0V9, Canada
            Tel: 204 787 2133
            Fax: 204 787 2190
            Email: hicksgg@cc.umanitoba.ca
            U3NeoSV1 gene trap. Tag generated by plasmid rescue. Additional
            sequence information and target gene cloning can be generated. ES
            cell line harboring insertion mutation of target gene is available.
            Sequence analysis available from
            http://140.193.242.7/esdb/public_search_frame.php?PST=PST3252-NR.Se
            q
            Class: Gene Trap.
FEATURES             Location/Qualifiers
     source          1. .35
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /strain="129 sv"
                     /db_xref="taxon:10090"
                     /clone="PST3252-NR.Seq"
                     /sex="Male"
                     /cell_type="Embryonic stem cell"
                     /cell_line="D3H (J1 subclone)"
                     /clone_lib="MICB1"
                     /note="Vector: U3NeoSV1"
ORIGIN

  Query Match            30.5%;  Score 13.4;  DB 9;  Length 35;
  Best Local Similarity  73.9%;  Pred. No. 6.4e+05;
  Matches   17;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          9 CGTTCCTTCTTAATAACCGGTCG 31
              | || |||||||||     |||||
Db          4 CTTTTCTTCTTAATTCTGGGTCG 26


RESULT 36
AZ642621
LOCUS       AZ642621                38 bp    DNA     linear   GSS 14-DEC-2000
DEFINITION  1M0505B13R Mouse 10kb plasmid UUGC1M library Mus musculus genomic
            clone UUGC1M0505B13 R, genomic survey sequence.
ACCESSION   AZ642621
VERSION     AZ642621.1  GI:11769410
KEYWORDS    GSS.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 38)
  AUTHORS   Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
            Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T.,
            Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von
            Niederhausern,A. and Wright,D.,Weiss,R.
  TITLE     Mouse whole genome scaffolding with paired end reads from 10kb
            plasmid inserts
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Robert B. Weiss
            University of Utah Genome Center
            University of Utah
            Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
            84112, USA
            Tel: 801 585 5606
            Fax: 801 585 7177
            Email: ddunn@genetics.utah.edu
            Insert Length: 10000   Std Error: 0.00
            Plate: 0505  row: B  column: 13
            Seq primer: CACACAGGAAACAGCTATGACC
            Class: plasmid ends
            High quality sequence stop: 38.
FEATURES             Location/Qualifiers
     source          1. .38
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="UUGC1M0505B13"
                     /sex="Male"
                     /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
                     /clone_lib="Mouse 10kb plasmid UUGC1M library"
                     /note="Vector: PWD42nv; Purified genomic DNA from M.
                     musculus C57BL/6J (male) was obtained  from the Jackson
                     Laboratory Mouse DNA Resource
                     (http://www.jax.org/resources/documents/dnares/). The DNA
                     was hydrodynamically sheared by repeated passage through a
                     0.005 inch orifice at constant velocity. The sheared DNA
                     was blunt end-repaired with T4 DNA polymerase and T4
                     polynucleotide kinase. Adaptor oligonucleotides were
                     ligated to the blunt ends in high  molar excess. The
                     adaptored DNA was purified and size-selected for a 9.5 to
                     10.5 kb range using preparative agarose gel
                     electrophoresis. Vector DNA was prepared from a derivative
                     of pWD42 (gi|4732114|gb|AF129072.1), a copy-number
                     inducible derivative of plasmid R1. The vector was ligated
                     with adaptors complementary to the insert adaptors and
                     purified. The sheared, adaptored mouse DNA was annealed to
                     adaptored vector DNA, and transformed into
                     chemically-competent E. coli XL10-Gold (Stratagene) cells
                     and selected for ampicillin resistance."
ORIGIN

  Query Match            30.5%;  Score 13.4;  DB 8;  Length 38;
  Best Local Similarity  64.5%;  Pred. No. 6.5e+05;
  Matches   20;  Conservative    0;  Mismatches   11;  Indels    0;  Gaps    0;

Qy          1 GCGGGTCCCGTTCCTTCTTAATAACCGGTCG 31
              | || |||| ||| | ||  || |||  | |
Db          5 GTGGCTCCCATTCATACTGCATCACCCCTGG 35


RESULT 37
AU010606/c
LOCUS       AU010606                39 bp    mRNA    linear   EST 31-JUL-1998
DEFINITION  AU010606 Schizosaccharomyces pombe late log phase cDNA
            Schizosaccharomyces pombe cDNA clone spc10127, mRNA sequence.
ACCESSION   AU010606
VERSION     AU010606.1  GI:3347286
KEYWORDS    EST.
SOURCE      Schizosaccharomyces pombe (fission yeast)
```

```
  ORGANISM  Schizosaccharomyces pombe
            Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
            Schizosaccharomycetales; Schizosaccharomycetaceae;
            Schizosaccharomyces.
REFERENCE   1  (bases 1 to 39)
  AUTHORS   Morimyo,M. and Mita,K.
  TITLE     Identification of expressed sequence tags of Schizosaccharomyces
            pombe
  JOURNAL   Unpublished (1998)
COMMENT     Contact: Mitsuoki Morimyo
            Genome Research Group
            National Institute of Radiological Sciences
            9-1, Anagawa-4-chome, Inage-ku, Chiba, Chiba 263-8555, Japan
            Email: morimyo@nirs.go.jp.
FEATURES             Location/Qualifiers
     source          1. .39
                     /organism="Schizosaccharomyces pombe"
                     /mol_type="mRNA"
                     /strain="972"
                     /db_xref="taxon:4896"
                     /clone="spc10127"
                     /sex="h minus"
                     /clone_lib="Schizosaccharomyces pombe late log phase cDNA"
                     /note="Vector: M13mp19; The cDNA library of
                     Schizosaccharomyces pombe was prepared by cloning cDNA
                     into the SmaI site of M13mp19 DNA and the direction of DNA
                     sequences was not always from 5' to 3'. The cDNA data of
                     Schizosaccharomyces pombe are available for searching on
                     the World Wide Web.  (URL, http://www.nirs.go.jp)"
ORIGIN

  Query Match             30.5%;  Score 13.4;  DB 1;  Length 39;
  Best Local Similarity   64.5%;  Pred. No. 6.5e+05;
  Matches   20;  Conservative    0;  Mismatches   11;  Indels    0;  Gaps    0;

Qy         14 CTTCTTAATAACCGGTCGCGGTTATTAAGAA 44
              | |  |||| || | | ||| |||  || ||
Db         32 CATAATAATTACTGATTGCGATTAAAAAAAA 2


RESULT 38
AU011134/c
LOCUS       AU011134                  39 bp    mRNA    linear   EST 03-AUG-1998
DEFINITION  AU011134 Schizosaccharomyces pombe late log phase cDNA
            Schizosaccharomyces pombe cDNA clone spc10805, mRNA sequence.
ACCESSION   AU011134
VERSION     AU011134.1  GI:3356043
KEYWORDS    EST.
SOURCE      Schizosaccharomyces pombe (fission yeast)
  ORGANISM  Schizosaccharomyces pombe
            Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
            Schizosaccharomycetales; Schizosaccharomycetaceae;
            Schizosaccharomyces.
REFERENCE   1  (bases 1 to 39)
  AUTHORS   Morimyo,M. and Mita,K.
  TITLE     Identification of expressed sequence tags of Schizosaccharomyces
            pombe
  JOURNAL   Unpublished (1998)
COMMENT     Contact: Mitsuoki Morimyo
            Genome Research Group
            National Institute of Radiological Sciences
            9-1, Anagawa-4-chome, Inage-ku, Chiba, Chiba 263-8555, Japan
            Email: morimyo@nirs.go.jp.
FEATURES             Location/Qualifiers
     source          1. .39
                     /organism="Schizosaccharomyces pombe"
                     /mol_type="mRNA"
                     /strain="972"
                     /db_xref="taxon:4896"
                     /clone="spc10805"
                     /sex="h minus"
                     /clone_lib="Schizosaccharomyces pombe late log phase cDNA"
                     /note="Vector: M13mp19; The cDNA library of
                     Schizosaccharomyces pombe was prepared by cloning cDNA
                     into the SmaI site of M13mp19 DNA and the direction of DNA
                     sequences was not always from 5' to 3'. The cDNA data of
                     Schizosaccharomyces pombe are available for searching on
                     the World Wide Web.  (URL, http://www.nirs.go.jp)"
ORIGIN

  Query Match             30.5%;  Score 13.4;  DB 1;  Length 39;
  Best Local Similarity   64.5%;  Pred. No. 6.5e+05;
  Matches   20;  Conservative    0;  Mismatches   11;  Indels    0;  Gaps    0;

Qy         14 CTTCTTAATAACCGGTCGCGGTTATTAAGAA 44
              | |  |||| || | | ||| |||  || ||
Db         32 CATAATAATTACTGATTGCGATTAAAAAAAA 2


RESULT 39
AU012105/c
LOCUS       AU012105                  39 bp    mRNA    linear   EST 03-AUG-1998
DEFINITION  AU012105 Schizosaccharomyces pombe late log phase cDNA
            Schizosaccharomyces pombe cDNA clone spc06348, mRNA sequence.
ACCESSION   AU012105
VERSION     AU012105.1  GI:3357014
KEYWORDS    EST.
SOURCE      Schizosaccharomyces pombe (fission yeast)
  ORGANISM  Schizosaccharomyces pombe
            Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
            Schizosaccharomycetales; Schizosaccharomycetaceae;
            Schizosaccharomyces.
REFERENCE   1  (bases 1 to 39)
  AUTHORS   Morimyo,M. and Mita,K.
  TITLE     Identification of expressed sequence tags of Schizosaccharomyces
            pombe
  JOURNAL   Unpublished (1998)
COMMENT     Contact: Mitsuoki Morimyo
            Genome Research Group
            National Institute of Radiological Sciences
            9-1, Anagawa-4-chome, Inage-ku, Chiba, Chiba 263-8555, Japan
            Email: morimyo@nirs.go.jp.
FEATURES             Location/Qualifiers
     source          1. .39
                     /organism="Schizosaccharomyces pombe"
                     /mol_type="mRNA"
                     /strain="972"
                     /db_xref="taxon:4896"
                     /clone="spc06348"
                     /sex="h minus"
                     /clone_lib="Schizosaccharomyces pombe late log phase cDNA"
                     /note="Vector: M13mp19; The cDNA library of
                     Schizosaccharomyces pombe was prepared by cloning cDNA
                     into the SmaI site of M13mp19 DNA and the direction of DNA
                     sequences was not always from 5' to 3'. The cDNA data of
                     Schizosaccharomyces pombe are available for searching on
                     the World Wide Web.  (URL, http://www.nirs.go.jp)"
ORIGIN

  Query Match             30.5%;  Score 13.4;  DB 1;  Length 39;
  Best Local Similarity   64.5%;  Pred. No. 6.5e+05;
  Matches   20;  Conservative    0;  Mismatches   11;  Indels    0;  Gaps    0;

Qy         14 CTTCTTAATAACCGGTCGCGGTTATTAAGAA 44
              | |  |||| || | | ||| |||  || ||
Db         32 CATAATAATTACTGATTGCGATTAAAAAAAA 2


RESULT 40
AU012381/c
LOCUS       AU012381                  39 bp    mRNA    linear   EST 03-AUG-1998
DEFINITION  AU012381 Schizosaccharomyces pombe late log phase cDNA
            Schizosaccharomyces pombe cDNA clone spc06701, mRNA sequence.
ACCESSION   AU012381
VERSION     AU012381.1  GI:3357290
KEYWORDS    EST.
```

```
SOURCE      Schizosaccharomyces pombe (fission yeast)
  ORGANISM  Schizosaccharomyces pombe
            Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
            Schizosaccharomycetales; Schizosaccharomycetaceae;
            Schizosaccharomyces.
REFERENCE   1  (bases 1 to 39)
  AUTHORS   Morimyo,M. and Mita,K.
  TITLE     Identification of expressed sequence tags of Schizosaccharomyces
            pombe
  JOURNAL   Unpublished (1998)
COMMENT     Contact: Mitsuoki Morimyo
            Genome Research Group
            National Institute of Radiological Sciences
            9-1, Anagawa-4-chome, Inage-ku, Chiba, Chiba 263-8555, Japan
            Email: morimyo@nirs.go.jp.
FEATURES             Location/Qualifiers
     source          1. .39
                     /organism="Schizosaccharomyces pombe"
                     /mol_type="mRNA"
                     /strain="972"
                     /db_xref="taxon:4896"
                     /clone="spc06701"
                     /sex="h minus"
                     /clone_lib="Schizosaccharomyces pombe late log phase cDNA"
                     /note="Vector: M13mp19; The cDNA library of
                     Schizosaccharomyces pombe was prepared by cloning cDNA
                     into the SmaI site of M13mp19 DNA and the direction of DNA
                     sequences was not always from 5' to 3'. The cDNA data of
                     Schizosaccharomyces pombe are available for searching on
                     the World Wide Web.  (URL, http://www.nirs.go.jp)"
ORIGIN

  Query Match             30.5%;  Score 13.4;  DB 1;  Length 39;
  Best Local Similarity   64.5%;  Pred. No. 6.5e+05;
  Matches   20;  Conservative    0;  Mismatches   11;  Indels    0;  Gaps    0;

Qy         14 CTTCTTAATAACCGGTCGCGGTTATTAAGAA 44
              | |  |||| || | | ||| |||  || ||
Db         32 CATAATAATTACTGATTGCGATTAAAAAAAA 2


Search completed: May 24, 2005, 13:16:40
Job time : 1861 secs
```